ASSIGNEE INFORMATION

The subject application (Attorney Docket No. DAS-101XC2; inventors **Scott Bintrim, Scott Bevan, Baolong Zhu,** and **Donald J. Merlo;** entitled **"Pesticidally Active Proteins and Polynucleotides Obtainable from *Paenibacillus* Species"**) has been assigned to:

Dow AgroSciences LLC
9330 Zionsville Road
Indianapolis, IN 46268-1054

DESCRIPTION

PESTICIDALLY ACTIVE PROTEINS AND POLYNUCLEOTIDES OBTAINABLE FROM *PAENIBACILLUS* SPECIES

Cross-Reference to Related Applications

[0001] This application claims priority to provisional application Serial No. 60/392,633, filed June 28, 2002, and to provisional application Serial No. 60/441,647, filed January 21, 2003.

Background of the Invention

[0002] Insects and other pests cost farmers billions of dollars annually in crop losses and in the expense of keeping these pests under control. The losses caused by insect pests in agricultural production environments include decreases in crop yield, reduced crop quality, and increased harvesting costs. Insect pests are also a burden to vegetable and fruit growers, to producers of ornamental flowers, and to home gardeners and homeowners.

[0003] Cultivation methods, such as crop rotation and the application of high levels of nitrogen fertilizers, have partially addressed problems caused by agricultural pests. However, economic demands on the utilization of farmland restrict the use of crop rotation. In addition, overwintering traits of some insects are disrupting crop rotations in some areas.

[0004] Thus, synthetic chemical insecticides are relied upon most heavily to achieve a sufficient level of control. However, the use of synthetic chemical insecticides can have several drawbacks. For example, the use of some of these chemicals can adversely affect many beneficial insects. Target insects have also developed resistance to some chemical pesticides. This has been partially alleviated by various resistance management strategies, but there is an increasing need for alternative pest control agents. Furthermore, very high populations of larvae, heavy rains, and improper calibration of insecticide application equipment can result in poor control. The improper use of insecticides raises environmental concerns such as contamination of soil and of both surface and underground water supplies. Residues can also remain on treated fruits, vegetables, and on other treated plants. Working with some insecticides can also pose hazards to the persons applying them. Therefore, synthetic chemical pesticides are being increasingly scrutinized for their potential toxic environmental consequences. Stringent new restrictions on

the use of pesticides and the elimination of some effective pesticides from the market place could limit economical and effective options for controlling damaging and costly pests.

[0005] Because of the problems associated with the use of synthetic chemical pesticides, there exists a clear need to limit the use of these agents and a need to identify alternative control agents. The replacement of synthetic chemical pesticides, or combination of these agents with biological pesticides, could reduce the levels of toxic chemicals in the environment.

[0006] Some biological pesticidal agents that are now being used with some success are derived from the soil microbe *Bacillus thuringiensis* (*B.t.*). The soil microbe *Bacillus thuringiensis* (*B.t.*) is a Gram-positive, spore-forming bacterium. Most strains of *B.t.* do not exhibit pesticidal activity. Some *B.t.* strains produce, and can be characterized by, parasporal crystalline protein inclusions. These inclusions often appear microscopically as distinctively shaped crystals. Some *B.t.* proteins are highly toxic to pests, such as insects, and are specific in their toxic activity. Certain insecticidal *B.t.* proteins are associated with the inclusions. These "δ-endotoxins" are different from exotoxins, which have a non-specific host range. Other species of *Bacillus* also produce pesticidal proteins.

[0007] Certain *Bacillus* toxin genes have been isolated and sequenced, and recombinant DNA-based products have been produced and approved for use. In addition, with the use of genetic engineering techniques, various approaches for delivering these toxins to agricultural environments are being perfected. These include the use of plants genetically engineered with toxin genes for insect resistance and the use of stabilized intact microbial cells as toxin delivery vehicles. Thus, isolated *Bacillus* toxin genes are becoming commercially valuable.

[0008] Commercial use of *B.t.* pesticides was initially restricted to targeting a narrow range of lepidopteran (caterpillar) pests. Preparations of the spores and crystals of *B. thuringiensis* subsp. *kurstaki* have been used for many years as commercial insecticides for lepidopteran pests. For example, *B. thuringiensis* var. *kurstaki* HD-1 produces a crystalline δ-endotoxin which is toxic to the larvae of a number of lepidopteran insects.

[0009] More recently, new subspecies of *B.t.* have been identified, and genes responsible for active δ-endotoxin proteins have been isolated. Höfte and Whiteley classified *B.t.* crystal protein genes into four major classes (Höfte, H., H.R. Whiteley [1989] *Microbiological Reviews* 52(2):242-255). The classes were *Cry*I (Lepidoptera-specific), *Cry*II (Lepidoptera- and Diptera-specific), *Cry*III (Coleoptera-specific), and *Cry*IV (Diptera-specific). The discovery of strains

specifically toxic to other pests has been reported. For example, *Cry*V and *Cry*VI were proposed to designate a class of toxin genes that are nematode-specific.

[0010] The Lepidopteran-specific *Cry*I crystal proteins, in their natural state, are approximately 130- to 140-kDa proteins, which accumulate in bipyramidal crystalline inclusions during the sporulation of *B. thuringiensis*. These proteins are protoxins which solubilize in the alkaline environment of the insect midgut and are proteolytically converted by crystal-associated or larval-midgut proteases into a toxic core fragment of 60 to 70 kDa. This activation can also be carried out *in vitro* with a variety of proteases. The toxic domain is localized in the N-terminal half of the protoxin. This was demonstrated for *Cry*IA(b) and *Cry*IC proteins through N-terminal amino acid sequencing of the trypsin-activated toxin. Höfte *et al.* 1989. Cleavage occurs on the C-terminal end of a conserved region called "Block 5," thus forming the C-terminus of the core toxin. A short, N-terminal protoxin segment can also be processed off. The N-terminal cleavage site is also highly conserved for *Cry*IA and *Cry*ID proteins, suggesting that for these proteins, the N terminus of the toxic fragment is localized at the same position. *Cry*IB, however, is different from the other *Cry*I proteins in this region. It was not known whether this protein is also processed at the N terminus. Höfte *et al.* 1989.

[0011] Deletion analysis of several *cry*I genes further confirmed that the 3' half of the protoxin is not required for toxic activity. One of the shortest reported toxic fragments was localized between codons 29 and 607 for *Cry*IAb. Further removal of four codons from the 3' end or eight codons from the 5' end completely abolished the toxic activity of the gene product. Similar observations were made for the *cry*IA(a) and *cry*IA(c) genes. Höfte *et al.* 1989.

[0012] The *cry*II genes encode 65-kDa proteins which form cuboidal inclusions in strains of several subspecies. These crystal proteins were previously designated "P2" proteins, as opposed to the 130-kDa P1 crystal proteins present in the same strains. Höfte *et al.* 1989.

[0013] A *cry*IIA gene was cloned from *B. thuringiensis* subsp. *kurstaki* HD-263 and expressed in *Bacillus megaterium*. Cells producing the *Cry*IIA protein were toxic for the lepidopteran species *Heliothis virescens* and *Lymantria dispar* as well as for larvae of the dipteran *Aedes aegypti*. Widner and Whitely (1989, *J. Bacteriol.* 171:965-974) cloned two related genes (*cry*IIA and *cry*IIB) from *B. thuringiensis* subsp. *kurstaki* HD-1. Both genes encode proteins of 633 amino acids with a predicted molecular mass of 71 kDa (slightly larger than the apparent molecular mass determined for the P2 proteins produced in *B. thuringiensis*). Although the *Cry*IIA and

*Cry*IIB proteins are highly homologous (~87% amino acid identity), they differ in their insecticidal spectra. *Cry*IIA is active against both a lepidopteran (*Manduca sexta*) and a dipteran (*Aedes aegypti*) species, whereas *cry*IIB is toxic only to the lepidopteran insect. Höfte *et al.* 1989. The *Cry*II toxins, as a group, tend to be relatively more conserved at the sequence level (>80% identical) than other groups. In contrast, there are many *Cry*I toxins, for example, including some that are less than 60% identical.

[0014] The 1989 nomenclature and classification scheme of Höfte and Whiteley for crystal proteins was based on both the deduced amino acid sequence and the host range of the toxin. That system was adapted to cover 14 different types of toxin genes which were divided into five major classes. The 1989 nomenclature scheme became unworkable as more and more genes were discovered that encoded proteins with varying spectrums of pesticidal activity. Thus, a revised nomenclature scheme was adopted, which is based solely on amino acid identity (Crickmore *et al.*, 1998, *Microbiology and Molecular Biology Reviews* 62:807-813). The mnemonic "*cry*" has been retained for all of the toxin genes except cytA and cytB, which remain a separate class. Roman numerals have been exchanged for Arabic numerals in the primary rank, and the parentheses in the tertiary rank have been removed. Many of the original names have been retained, with the noted exceptions, although a number have been reclassified. There are now at least 37 primary classes of *Cry* proteins, and two primary classes of *cyt* toxins. Other types of toxins, such as those of WO 98/18932 and WO 97/40162, have also been discovered from *B. thuringiensis.*

[0015] There are some obstacles to the successful agricultural use of *Bacillus* (and other biological) pesticidal proteins. Certain insects can be refractory to the effects of *Bacillus* toxins. Insects such as boll weevils, black cutworm, and *Helicoverpa zea*, as well as adult insects of most species, heretofore have demonstrated no significant sensitivity to many *B.t.* δ-endotoxins.

[0016] Another potential obstacle is the development of resistance to *B.t.* toxins by insects. *B.t.* protein toxins were initially formulated as sprayable insect control agents. A more recent application of *B.t.* technology has been to isolate and transform plants with genes that encode these toxins. Transgenic plants subsequently produce the toxins, thereby providing insect control. *See* U.S. Patent Nos. 5,380,831; 5,567,600; and 5,567,862 to Mycogen Corporation. Transgenic *B.t.* plants are quite efficacious, and usage is predicted to be high in some crops and areas. This has caused some concern that resistance management issues may arise more quickly than with

traditional sprayable applications. While a number of insects have been selected for resistance to *B.t.* toxins in the laboratory, only the diamondback moth (*Plutella xylostella*) has demonstrated resistance in a field setting (Ferre, J. and Van Rie, J., *Annu. Rev. Entomol.* 47:501-533, 2002).

[0017] Resistance management strategies in *B.t.* transgene plant technology have become of great interest (for example, as in a natural bacterium, multiple diverse toxins can be exposed on the same plant, thereby greatly reducing the chance that an insect that might be resistant to one toxin would survive to spread the resistance). Several strategies have been suggested for preserving the ability to effectively use *B. thuringiensis* toxins. These strategies include high dose with refuge, and alternation with, or co-deployment of, different toxins (McGaughey *et al.* (1998), "*B.t.* Resistance Management," *Nature Biotechnol* 16:144-146).

[0018] Thus, there remains a great need for developing additional genes that can be expressed in plants in order to effectively control various insects. In addition to continually trying to discover new *B.t.* toxins, it would be quite desirable to discover other bacterial sources (distinct from *B.t.*) that produce toxins that could be used in transgenic plant strategies, or that could be combined with *B.t.*s to produce insect-controlling transgenic plants.

[0019] The recent efforts to clone insecticidal toxin genes from the *Photorhabdus/Xenorhabdus* group of bacteria present potential alternatives to toxins derived from *B. thuringiensis*. It has been known in the art that bacteria of the genus *Xenorhabdus* are symbiotically associated with the *Steinernema* nematode. Unfortunately, as reported in a number of articles, the bacteria only had pesticidal activity when injected into insect larvae and did not exhibit biological activity when delivered orally.

[0020] It has been difficult to effectively exploit the insecticidal properties of the nematode or its bacterial symbiont. Thus, it would be quite desirable to discover proteinaceous agents from *Xenorhabdus* bacteria that have oral activity so that the products produced therefrom could be formulated as a sprayable insecticide, or the bacterial genes encoding said proteinaceous agents could be isolated and used in the production of transgenic plants. WO 95/00647 relates to the use of *Xenorhabdus* protein toxin to control insects, but it does not recognize orally active toxins. WO 98/08388 relates to orally administered pesticidal agents from *Xenorhabdus*. U.S. Patent No. 6,048,838 relates to protein toxins/toxin complexes, having oral activity, obtainable from *Xenorhabdus* species and strains.

[0021] *Photorhabdus* and *Xenorhabdus* spp. are Gram-negative bacteria that entomopathogenically and symbiotically associate with soil nematodes. These bacteria are found in the gut of entomopathogenic nematodes that invade and kill insects. When the nematode invades an insect host, the bacteria are released into the insect haemocoel (the open circulatory system), and both the bacteria and the nematode undergo multiple rounds of replication; the insect host typically dies. These bacteria can be cultured away from their nematode hosts. For a more detailed discussion of these bacteria, *see* Forst and Nealson, 60 *Microbiol. Rev.* 1 (1996), pp. 21-43.

[0022] The genus *Xenorhabdus* is taxonomically defined as a member of the Family Enterobacteriaceae, although it has certain traits atypical of this family. For example, strains of this genus are typically nitrate reduction negative and catalase negative. *Xenorhabdus* has only recently been subdivided to create a second genus, *Photorhabdus,* which is comprised of the single species *Photorhabdus luminescens* (previously *Xenorhabdus luminescens*) (Boemare *et al.*, 1993 *Int. J. Syst. Bacteriol.* 43, 249-255). This differentiation is based on several distinguishing characteristics easily identifiable by the skilled artisan. These differences include the following: DNA-DNA characterization studies; phenotypic presence (*Photorhabdus*) or absence (*Xenorhabdus*) of catalase activity; presence (*Photorhabdus*) or absence (*Xenorhabdus*) of bioluminescence; the Family of the nematode host in that *Xenorhabdus* is found in *Steinernematidae* and *Photorhabdus* is found in *Heterorhabditidae*); as well as comparative, cellular fatty-acid analyses (Janse *et al.* 1990, *Lett. Appl. Microbiol.* 10, 131-135; Suzuki *et al.* 1990, *J. Gen. Appl. Microbiol.*, 36, 393-401). In addition, recent molecular studies focused on sequence (Rainey *et al.* 1995, *Int. J. Syst. Bacteriol.*, 45, 379-381) and restriction analysis (Brunel *et al.*, 1997, *App. Environ. Micro.*, 63, 574-580) of 16S rRNA genes also support the separation of these two genera.

[0023] The expected traits for *Xenorhabdus* are the following: Gram stain negative rods, white to yellow/brown colony pigmentation, presence of inclusion bodies, absence of catalase, inability to reduce nitrate, absence of bioluminescence, ability to uptake dye from medium, positive gelatin hydrolysis, growth on Enterobacteriaceae selective media, growth temperature below 37° C, survival under anaerobic conditions, and motility.

[0024] Currently, the bacterial genus *Xenorhabdus* is comprised of four recognized species, *Xenorhabdus nematophilus*, *Xenorhabdus poinarii*, *Xenorhabdus bovienii* and *Xenorhabdus*

beddingii (Brunel *et al.*, 1997, *App. Environ. Micro.*, 63, 574-580). A variety of related strains have been described in the literature (*e.g.*, Akhurst and Boemare 1988 *J. Gen. Microbiol.*, 134, 1835-1845; Boemare *et al.* 1993 *Int. J. Syst. Bacteriol.* 43, pp. 249-255; Putz *et al.* 1990, *Appl. Environ. Microbiol.*, 56, 181-186, Brunel *et al.*, 1997, *App. Environ. Micro.*, 63, 574-580, Rainey *et al.* 1995, *Int. J. Syst. Bacteriol.*, 45, 379-381).

[0025] *Xenorhabdus* and *Photorhabdus* bacteria secrete a wide variety of substances into the culture medium; these secretions include lipases, proteases, antibiotics and lipopolysaccharides. Purification of different protease fractions has clearly demonstrated that they are not involved in the oral toxic activity of *P. luminescens* culture medium (which has been subsequently determined to reside with the Tc proteins only). Several of these substances have previously been implicated in insect toxicity but until recently no insecticidal genes had been cloned. However, protease purification and separation will also facilitate an examination of their putative role in, for example, inhibiting antibacterial proteins such as cecropin. R.H. ffrench-Constant and Bowen, *Current Opinions in Micriobiology,* 1999, 12:284-288. *See* R.H. ffrench-Constant *et al.* 66 *AEM* No. 8, pp. 3310-3329 (Aug. 2000), for a review of various factors involved in *Photorhabdus* virulence of insects.

[0026] There has been substantial progress in the cloning of genes encoding insecticidal toxins from both *Photorhabdus luminescens* and *Xenorhabdus nematophilus*. Toxin-complex encoding genes from *P. luminescens* were examined first. *See, e.g.,* WO 98/08932. "Parallel" genes were more recently cloned from *X. nematophilus.* Morgan *et al., Applied and Environmental Microbiology* 2001, 67:2062-69.

[0027] Four different toxin complexes (TCs)—Tca, Tcb, Tcc and Tcd—have been identified in *Photorhabdus* spp. Each of these toxin complexes resolves as either a single or dimeric species on a native agarose gel but resolution on a denaturing gel reveals that each complex consists of a range of species between 25-280 kDa. The ORFs that encode the TCs from *Photorhabdus,* together with protease cleavage sites (vertical arrows), are illustrated in **Figure 1.** *See also* R.H. ffrench-Constant and Bowen, 57 *Cell. Mol. Life Sci.* 828-833 (2000).

[0028] Genomic libraries of *P. luminescens* were screened with DNA probes and with monoclonal and/or polyclonal antibodies raised against the toxins. Four *tc* loci were cloned: *tca, tcb, tcc* and *tcd*. The *tca* locus is a putative operon of three open reading frames (ORFs), *tcaA*, *tcaB*, and *tcaC* transcribed from the same DNA strand, with a smaller terminal ORF (*tcaZ*)

transcribed in the opposite direction. The *tcc* locus also is comprised of three ORFs putatively transcribed in the same direction (*tccA*, *tccB*, and *tccC*). The *tcb* locus is a single large ORF (*tcbA*), and the *tcd* locus is composed of two ORFs (*tcdA* and *tcdB*); *tcbA* and *tcdA*, each about 7.5 kb, encode large insect toxins. TcdB has some homology to TcaC. Many of these gene products were determined to be cleaved by proteases. For example, both TcbA and TcdA are cleaved into three fragments termed i, ii and iii (*e.g.* TcbAi, TcbAii and TcbAiii). Products of the *tca* and *tcc* ORFs are also cleaved. *See* **Figure 1.** *See also* R.H. ffrench-Constant and D.J. Bowen, *Current Opinions in Microbiology,* 1999, 12:284-288.

[0029] Bioassays of the Tca toxin complexes revealed them to be highly toxic to first instar tomato hornworms (*Manduca sexta*) when given orally (LD_{50} of 875 ng per square centimeter of artificial diet). R.H. ffrench-Constant and Bowen 1999. Feeding was inhibited at Tca doses as low as 40 ng/cm^2. Given the high predicted molecular weight of Tca, on a molar basis, *P. luminescens* toxins are highly active and relatively few molecules appear to be necessary to exert a toxic effect. R.H. ffrench-Constant and Bowen, *Current Opinions in Micriobiology,* 1999, 12:284-288.

[0030] None of the four loci showed overall similarity to any sequences of known function in GenBank. Regions of sequence similarity raised some suggestion that these proteins (TcaC and TccA) may overcome insect immunity by attacking insect hemocytes. R.H. ffrench-Constant and Bowen, *Current Opinions in Microbiology,* 1999, 12:284-288.

[0031] TcaB, TcbA, and TcdA all show amino acid conservation (~50% identity), compared with each other, immediately around their predicted protease cleavage sites. This conservation between three different TC proteins suggests that they may all be processed by the same or similar proteases. TcbA and TcdA also share ~50% identity overall, as well as a similar predicted pattern of both carboxy- and amino-terminal cleavage. It was postulated that these proteins might thus be homologs of one another. Furthermore, the similar, large size of TcbA and TcdA, and also the fact that both toxins appear to act on the gut of the insect, may suggest similar modes of action. R.H. ffrench-Constant and Bowen, *Current Opinions in Microbiology,* 1999, 12:284-288.

[0032] Deletion/knock-out studies suggest that products of the *tca* and *tcd* loci account for the majority of oral toxicity to lepidopterans. Deletion of either of the *tca* or *tcd* genes greatly reduced oral activity against *Manduca sexta.* That is, products of the *tca* and *tcd* loci are oral

lepidopteran toxins on their own; their combined effect contributed most of the secreted oral activity. R.H. ffrench-Constant and D.J. Bowen, 57 *Cell. Mol. Life. Sci.* 831 (2000). Interestingly, deletion of either of the *tcb* or *tcc* loci alone also reduces mortality, suggesting that there may be complex interactions among the different gene products. Thus, products of the *tca* locus may enhance the toxicity of *tcd* products. Alternatively, *tcd* products may modulate the toxicity of *tca* products and possibly other complexes. Noting that the above relates to oral activity against a single insect species, *tcb* or *tcc* loci may produce toxins that are more active against other groups of insects (or active via injection directly into the insect haemocoel—the normal route of delivery when secreted by the bacteria *in vivo*). R.H. ffrench-Constant and Bowen, *Current Opinions in Microbiology,* 1999, 12:284-288.

[0033] WO 01/11029 discloses nucleotide sequences that encode TcdA and TcbA and have base compositions that have been altered from that of the native genes to make them more similar to plant genes. Also disclosed are transgenic plants that express Toxin A and Toxin B.

[0034] Of the separate toxins isolated from *Photorhabdus luminescens* (W-14), those designated Toxin A and Toxin B have been the subject of focused investigation for their activity against target insect species of interest (*e.g.,* corn rootworm). Toxin A is comprised of two different subunits. The native gene tcdA encodes protoxin TcdA. As determined by mass spectrometry, TcdA is processed by one or more proteases to provide Toxin A. More specifically, TcdA is an approximately 282.9 kDa protein (2516 aa) that is processed to provide TcdAi (the first 88 amino acids), TcdAii (the next 1849 aa; an approximately 208.2 kDa protein encoded by nucleotides 265-5811 of *tcdA*), and TcdAiii, an approximately 63.5 kDa (579 aa) protein (encoded by nucleotides 5812-7551 of *tcdA*). TcdAii and TcdAiii appear to assemble into a dimer (perhaps aided by TcdAi), and the dimers assemble into a tetramer of four dimers. Toxin B is similarly derived from TcbA.

[0035] While the exact molecular interactions of the TC proteins with each other, and their mechanism(s) of action, are not currently understood, it is known, for example, that the Tca toxin complex of *Photorhabdus* is toxic to *Manduca sexta.* In addition, some TC proteins are known to have "stand alone" insecticidal activity, while other TC proteins are known to potentiate or enhance the activity of the stand-alone toxins. It is known that the TcdA protein is active, alone, against *Manduca sexta,* but that TcdB and TccC, together, can be used to enhance the activity of TcdA. Waterfield, N. *et al., Appl. Environ. Microbiol.* 2001, 67:5017-5024. TcbA (there is

only one Tcb protein) is another stand-alone toxin from *Photorhabdus*. The activity of this toxin (TcbA) can also be enhanced by TcdB together with TccC-like proteins.

[0036] U.S. Patent Application 20020078478 provides nucleotide sequences for two potentiator genes, tcdB2 and tccC2, from the tcd genomic region of *Photorhabdus luminescens* W-14. It is shown therein that coexpression of tcdB and tccC1 with tcdA results in enhanced levels of oral insect toxicity compared to that obtained when tcdA is expressed alone. Coexpression of tcdB and tccC1 with tcdA or tcbA provide enhanced oral insect activity.

[0037] As indicated in the chart below, TccA has some level of homology with the N terminus of TcdA, and TccB has some level of homology with the C terminus of TcdA. TccA and TccB are much less active on certain test insects than is TcdA. TccA and TccB from Photorhabdus strain W-14 are called "Toxin D." "Toxin A" (TcdA), "Toxin B" (TcbA), and "Toxin C" (TcaA and TcaB) are also indicated below. Furthermore, TcaA has some level of homology with TccA and likewise with the N terminus of TcdA. Still further, TcaB has some level of homology with TccB and likewise with the N terminus of TcdA. TccA and TcaA are of a similar size, as are TccB and TcaB. TcdB has a significant level of similarity (both in sequence and size) to TcaC.

Photorhabdus	*Photorhabdus strain W14 nomenclature*	Some homology to:
TcaA	Toxin C	TccA
TcaB		TccB
TcaC		TcdB
TcbA	Toxin B	
TccA	Toxin D	TcdA N terminus
TccB		TcdA C terminus
TccC		
TcdA	Toxin A	TccA + TccB
TcdB		TcaC

[0038] The insect midgut epithelium contains both columnar (structural) and goblet (secretory) cells. Ingestion of *tca* products by *M. sexta* leads to apical swelling and blebbing of large cytoplasmic vesicles by the columnar cells, leading to the eventual extrusion of cell nuclei in vesicles into the gut lumen. Goblet cells are also apparently affected in the same fashion. Products of *tca* act on the insect midgut following either oral delivery or injection. R.H. ffrench-Constant and D.J. Bowen, *Current Opinions in Microbiology,* 1999, 12:284-288. Purified *tca* products have shown oral toxicity against *Manduca sexta* (LD_{50} of 875 ng/cm^2). R.H. ffrench-Constant and D.J. Bowen, 57 *Cell. Mol. Life Sci.* 828-833 (2000).

[0039] WO 99/42589 and U.S. Patent No. 6,281,413 disclose TC-like ORFs from *Photorhabdus luminescens*. WO 00/30453 and WO 00/42855 disclose TC-like proteins from *Xenorhabdus*. WO 99/03328 and WO 99/54472 (and U.S. Patent Nos. 6,174,860 and 6,277,823) relate to other toxins from *Xenorhabdus* and *Photorhabdus*.

[0040] Relatively recent cloning efforts in *Xenorhabdus nematophilus* also appear to have identified novel insecticidal toxin genes with homology to the *P. luminescens tc* loci. *See, e.g.*, WO 98/08388 and Morgan *et al., Applied and Environmental Microbiology* 2001, 67:2062-69. In R.H. ffrench-Constant and D.J. Bowen, *Current Opinions in Microbiology,* 1999, 12:284-288, cosmid clones were screened directly for oral toxicity to another lepidopteran, *Pieris brassicae*. One orally toxic cosmid clone was sequenced. Analysis of the sequence in that cosmid suggested that there are five different ORF's with similarity to *Photorhabdus tc* genes; *orf2* and *orf5* both have some level of sequence relatedness to both *tcbA* and *tcdA,* whereas *orf1* is similar to *tccB, orf3* is similar to *tccC* and *orf4* is similar to *tcaC*. Importantly, a number of these predicted ORFs also share the putative cleavage site documented in *P. luminescens*, suggesting that active toxins may also be protealytically processed.

[0041] There are five typical *Xenorhabdus* TC proteins: XptA1, XptA2, XptB1, XptC1, and XptD1. XptA1 is a "stand-alone" toxin. XptA2 is another TC protein from *Xenorhabdus* that has stand-alone toxin activity. *See* GENBANK Accession No. AJ308438 for sequences from *Xenorhabdus nematophilus*. XptB1 and XptC1 are the *Xenorhabdus* potentiators that can enhance the activity of either (or both) of the XptA toxins. XptD1 has some level of homology with TccB. XptC1 has some level of similarity to TcaC. The XptA2 protein of *Xenorhabdus* has some degree of similarity to the TcdA protein. XptB1 has some level of similarity to TccC.

[0042] The finding of somewhat similar, toxin-encoding loci in these two different bacteria is interesting in terms of the possible origins of these virulence genes. The *X. nematophilus* cosmid also appears to contain transposase-like sequences whose presence may suggest that these loci can be transferred horizontally between different strains or species of bacteria. A range of such transfer events may also explain the apparently different genomic organization of the *tc* operons in the two different bacteria. Further, only a subset of *X. nematophilus* and *P. luminescens* strains appear toxic to *M. sexta*, suggesting either that different strains lack the *tc* genes or that they carry a different *tc* gene compliment. Detailed analysis of both a strain and toxin phylogeny within, and between, these bacterial species should help clarify the likely origin of the toxin genes and

how they are maintained in different bacterial populations. R.H. ffrench-Constant and Bowen, *Current Opinions in Microbiology,* 1999, 12:284-288.

[0043] TC proteins and genes have more recently been described from other insect-associated bacteria such as *Serratia entomophila*, an insect pathogen. Waterfield *et al.*, *TRENDS in Microbiology,* Vol. 9, No. 4, April 2001.

[0044] In summary, toxin complex proteins from *P. luminescens* and *X. nematophilus* appear to have little homology to previously identified bacterial toxins and should provide useful alternatives to toxins derived from *B. thuringiensis*. Although they have similar toxic effects on the insect midgut to other orally active toxins, their precise mode of action remains obscure. Future work could clarify their mechanism of action.

[0045] Although some *Xenorhabdus* TC proteins were found to "correspond" (have a similar function and some level of sequence homology) to some of the *Photorhabdus* TC proteins, a given *Photorhabdus* protein shares only about 40% sequence identity with the "corresponding" *Xenorhabdus* protein. This is illustrated below for four "stand-alone" toxins:

	Identity to P.l. W-14 TcbA	Identity to P.l. W-14 TcdA
Xwi XptA1	44%	46%
Xwi XptA2	41%	41%

(For a more complete review, *see, e.g.,* Morgan *et al.*, "Sequence Analysis of Insecticidal Genes from *Xenorhabdus nematophiles* PMFI296," Vol. 67, *Applied and Environmental Microbiology,* May 2001, pp. 2062-2069.)

[0046] Bacteria of the genus *Paenibacillus* are distinguishable from other bacteria by distinctive rRNA and phenotypic characteristics (C. Ash *et al:* (1993), "Molecular identification of rRNA group 3 bacilli (Ash, Farrow, Wallbanks and Collins) using a PCR probe test: Proposal for the creation of a new genus *Paenibacillus*," *Antonie Van Leeuwenhoek* 64:253-260). Comparative 16S rRNA sequence analysis demonstrated that the genus *Bacillus* consisted of at least five phyletic lines. Ribosomal RNA group 3 bacilli (of Ash, Farrow, Wallbanks, and Collins (1991), comprising *Bacillus polymyxa* and close relatives), is phylogenetically so removed from *Bacillus subtilis* (the type species of the genus and other aerobic, endospore-forming bacilli) that they were reclassified as a new genus, *Paenibacillus.*

[0047] Some species in this genus were known to be pathogenic to honeybees (*Paenibacillus Larvae*) and scarab beetle grubs (*P. popilliae* and *P. lentimorbus*). Some other *Paenibacillus* species that have been found to be associated with honeybees, but they are non-pathogens. At least 18 additional species are known in this genus, including *P. thiaminolyticus;* they have no known insect association (Shida *et al.*, 1997; Pettersson *et al.*, 1999). Scarabs (coleopterans) are serious pests of turf, nurseries, and food crops throughout North America, and are of quarantine concern. *See* U.S. Department of Agriculture, Agricultural Research Service website.

[0048] *P. larvae, P. popilliae,* and *P. lentimorbus* are considered obligate insect pathogens involved with milky disease of scarab beetles (D.P. Stahly *et al.* (1992), "The genus *Bacillus:* insect pathogens," p. 1697-1745, *In* A. Balows *et al.*, ed., *The Procaryotes,* 2nd Ed., Vol. 2, Springer-Verlag, New York, NY). These three *Paenibacillus* species are characteristically slow-growing, fastidious organisms that cause disease by an invasive process in which the bacteria cross the midgut and proliferate to high numbers in the hemolymph and other tissues. For all three species, some general indications of protein involvement in insect pathogenicity have been proposed; however, no specific role for a specific protein has been demonstrated. Stahly *et al.* concluded for *P. larvae* that a question of the involvement of a toxin is an open one, and that the precise cause of death in milky disease (of beetles) is not understood.

[0049] A beetle (coleopteran) toxin, *Cry*18, has been identified in strains of *P. popilliae* and *P. lentimorbus. Cry*18 has about 40% identity to *Cry*2 proteins (Zhang *et al.,* 1997; Harrison *et al.,* 2000). While Zhang *et al.* (1997) speculate that *Cry*18 attacks the midgut to facilitate entry of vegetative cells to the hemocoel, Harrison *et al.* note that there is no direct evidence for this role and further state that "the role, if any, of the paraspore protein in milky disease is unknown." J. Zhang *et al.* (1997), "Cloning and Analysis of the First *cry* Gene from *Bacillus popilliae,*" *J. Bacteriol.* 179:4336-4341; H. Harrison *et al.* (2000), "*Paenibacillus* Associated with Milky Disease in Central and South American Scarabs," *J. Invertebr. Pathol.* 76(3):169-175.

[0050] Stahly *et al.*, Zhang *et al.*, and Harrison *et al.* all point to the contrast in evidence for the role of crystal proteins of *B. thuringiensis* in intoxication of insects (where the high frequency of insect symptoms can be explained by the properties of the specific crystal proteins), versus the case of *Paenibacillus* and milky disease (where there is no such tie to the effects of a specific toxin).

[0051] Thus, while some species of *Paenibacillus* were known to be pathogenic to certain coleopterans and some associated with honeybees, no strain of *Paenibacillus* was heretofore known to be toxic to lepidopterans. Likewise, TC proteins and lepidopteran-toxic *Cry* proteins have never been reported in *Paenibacillus*.

Brief Summary of the Invention

[0052] This is the first known disclosure of *Paenibacillus* protein toxins having activity against lepidopteran pests. Some species of *Paenibacillus* were known to be insecticidal, but they had activity against grubs/beetles/coleopterans. There have been no known reports of a *Paenibacillus* species or strain having toxicity to lepidopterans. Thus, the subject invention relates generally to *Paenibacillus* species that have activity against lepidopterans, and to screening *Paenibacillus* spp., proteins therefrom, and libraries of clones therefrom for activity against lepidopterans.

[0053] More specifically, the subject invention initially stemmed from a discovery of a novel strain of *Paenibacillus* referred to herein as DAS1529. This was a surprising discovery for a variety of reasons. This strain produces a unique, lepidopteran-toxic *Cry* protein. This strain, as well as DB482, produce unique, toxin complex (TC)-like proteins (having some similarity to *Xenorhabdus/Photorhabdus* TCs). *Paenibacillus* isolate DB482 and toxins obtainable therefrom are highly preferred, and all are within the scope of the subject invention.

[0054] This is the first known report of *Paenibacillus* having TC-like proteins. Thus, the subject invention relates to methods of screening *Paenibacillus* spp. for TC-like genes and proteins. *Paenibacillus* TC proteins of the subject invention are shown herein to be useful to enhance or potentiate the activity of a "stand-alone" *Xenorhabdus* toxin protein, for example. TC-like genes identified herein were not heretofore known to exist in the genus *Paenibacillus*. This discovery broadens the scope of organisms (bacterial genera) in which TC-like genes have been found. Thus, the subject invention generally relates to TC-like proteins obtainable from *Paenibacillus* species, to methods of screening *Paenibacillus* species for such proteins, and the like. One example is *Paenibacillus apairius,* which was also found to produce TC-like proteins.

[0055] While the subject TC-like proteins have some sequence relatedness to, and characteristics in common with, TC proteins of *Xenorhabdus* and *Photorabdus*, the sequences of the subject TC-like proteins are very different from previously known TC proteins. Thus, the subject application

provides new classes of TC-like proteins and genes that encode these proteins, which are obtainable from bacteria in the genera *Paenibacillus, Photorhabdus, Xenorhabdus,* and the like.

[0056] Another surprising feature of the DAS1529 strain is that it produces a unique, *B.t.*-like *Cry* protein that is toxic to lepidopterans. The subject *Cry* toxin is compressed/short and appears to lack a typical protoxin portion in its wild-type state. Thus, the subject invention generally relates to screening *Paenibacillus* isolates for lepidopteran-toxic *Cry* proteins. The subject invention also relates to methods of screening *Paenibacillus* spp. and *B. thuringiensis,* for example, for this new class of *Cry* genes and proteins.

[0057] The DAS1529 strain is the first known example of a natural bacterium that produces both a *Cry*-like toxin and TC-like proteins. Further surprising is that this is the first known example of a *cry* toxin gene being closely associated with (in genetic proximity to) TC protein genes. These pioneering observations have broad implications and thus enable one skilled in the art to screen appropriate species of bacteria for these types of unique operons and for these types of further components of known operons. Such techniques are within the scope of the subject invention.

[0058] A further aspect of the subject invention stems from the surprising discovery that the DAS1529 strain also produces a soluble insect toxin that was found to be very similar to a thiaminase. It was surprising that the *Paenibacillus* thiaminase protein was found to have insecticidal activity. While this type of protein was known, it was in no way expected in the art that this enzyme would have exhibited toxin-like activity against insects/insect-like pests. Thus, the subject invention also relates to methods of screening *Paenibacillus* and others for insecticidal thiaminase genes and proteins, and to the use of these genes and proteins for controlling insects and like pests.

[0059] Other objects, advantages, and features of the subject invention will be apparent to one skilled in the art having the benefit of the subject disclosure.

Brief Description of the Figures

[0060] **Figure 1** shows the TC operons from *Photorhabdus*.

[0061] **Figure 2** shows a diagram of the DNA from DAS1529 inserted into the "SB12" clone that exhibited pesticidal activity, with open reading frames identified with block and line arrows.

[0062] **Figure 3** shows partial sequence alignments for SEQ ID NO:17 and thiaminase I from *Bacillus thiaminolyticus* (Campobasso *et al.,* 1998) or AAC44156.

[0063] **Figure 4** shows test results of purified thiaminase from DAS1529 on CEW.

[0064] **Figure 5** shows ORF3-ORF6 in pEt101D.

[0065] **Figure 6** shows *Cry*1529 (ORF 7) against tobacco bud worm (TBW).

[0066] **Figure 7** shows a comparison/alignment of SEQ ID NO:9 to SEQ ID NO:5 ($tcaB_2$ to $tcaB_1$); the brackets show the ORF2 junction.

[0067] **Figure 8** shows a phylogenetic tree of DAS1529 ORF7 (*Cry*1529) compared to other *Cry* proteins.

[0068] **Figures 9 and 10** show results of trypsin digestion of wild-type and modified *Cry*1529 proteins.

[0069] **Figures 11A and 11B** show sequence alignments for tcaA primer design.

[0070] **Figures 12A-D** show sequence alignments for tcaB primer design.

[0071] **Figures 13A and 13B** show sequence alignments for tcaC primer design.

[0072] **Figures 14A and 14B** show sequence alignments for tccC primer design.

Brief Description of the Sequences

[0073] **SEQ ID NO:1** is the nucleic acid sequence of the entire insert of SB12.

[0074] **SEQ ID NO:2** is the nucleic acid sequence of ORF1, which encodes a tcaA-like protein (gene *tcaA1,* source organism *Paenibacillus* strain IDAS 1529, gene designation *tcaA1-1529*).

[0075] **SEQ ID NO:3** is the amino acid sequence encoded by ORF1.

[0076] **SEQ ID NO:4** is the nucleic acid sequence of ORF2, with an IS element removed, which encodes a tcaB-like protein (gene *tcaB1,* source organism *Paenibacillus* strain IDAS 1529, gene designation *tcaB1-1529*).

[0077] **SEQ ID NO:5** is the amino acid sequence encoded by ORF2.

[0078] **SEQ ID NO:6** is the nucleic acid sequence of ORF3, which encodes a tcaA-like protein (gene *tcaA2,* source organism *Paenibacillus* strain IDAS 1529, gene designation *tcaA2-1529*).

[0079] **SEQ ID NO:7** is the amino acid sequence encoded by ORF3.

[0080] **SEQ ID NO:8** is the nucleic acid sequence of ORF4, which encodes a tcaB-like protein (gene *tcaB2,* source organism *Paenibacillus* strain IDAS 1529, gene designation *tcaB2-1529*).

[0081] **SEQ ID NO:9** is the amino acid sequence encoded by ORF4.

[0082] **SEQ ID NO:10** is the nucleic acid sequence of ORF5, which encodes a tcaC-like protein (gene *tcaC,* source organism *Paenibacillus* strain IDAS 1529, gene designation *tcaC-1529*).

[0083] **SEQ ID NO:11** is the amino acid sequence encoded by ORF5.

[0084] **SEQ ID NO:12** is the nucleic acid sequence of ORF6, which encodes a tccC-like protein.

[0085] **SEQ ID NO:13** is the amino acid sequence encoded by ORF6.

[0086] **SEQ ID NO:14** is the nucleic acid sequence of ORF7, which encodes a *Cry*-like protein.

[0087] **SEQ ID NO:15** is the amino acid sequence encoded by ORF7.

[0088] **SEQ ID NO:16** is the partial nucleic acid sequence of the 16S rDNA of DAS1529 used for taxonomic placement.

[0089] **SEQ ID NO:17** is the N-terminal amino acid sequence for the purified toxin from the broth fraction from DAS1529.

[0090] **SEQ ID NO:18** is the amino acid sequence of thiaminase I from *Bacillus thiaminolyticus* (Campobasso *et al., J. Biochem.* 37(45):15981-15989 (1998)).

[0091] **SEQ ID NO:19** is an alternate amino acid sequence encoded by ORF6 protein (gene *tccC,* source organism *Paenibacillus* strain IDAS 1529, gene designation *tccC-1529*).

[0092] **SEQ ID NO:20** is gene *xptC1,* source organism *Xenorhabdus* strain Xwi, gene designation xptC1-Xwi.

[0093] **SEQ ID NO:21** is gene *xptB1,* source organism *Xenorhabdus* strain Xwi, gene designation xptB1-Xwi.

[0094] **SEQ ID NO:22** is primer SB101.

[0095] **SEQ ID NO:23** is primer SB102.

[0096] **SEQ ID NO:24** is primer SB103.

[0097] **SEQ ID NO:25** is primer SB104.

[0098] **SEQ ID NO:26** is primer SB105.

[0099] **SEQ ID NO:27** is primer SB106.

[00100] **SEQ ID NO:28** is primer SB212.

[00101] **SEQ ID NO:29** is primer SB213.

[00102] **SEQ ID NO:30** is primer SB215.

[00103] **SEQ ID NO:31** is primer SB217.

[00104] **SEQ ID NO:32** is a nucleotide sequence from a tcaA-like gene from *Paenibacillus apairius* strain DB482.

[00105] **SEQ ID NO:33** is an amino acid sequence from a TcaA-like protein from *Paenibacillus apairius* strain DB482.

[00106] **SEQ ID NO:34** is a nucleotide sequence from a tcaB-like gene from *Paenibacillus apairius* strain DB482.

[00107] **SEQ ID NO:35** is a nucleotide sequence from a tcaB-like gene from *Paenibacillus apairius* strain DB482.

[00108] **SEQ ID NO:36** is an amino acid sequence from a TcaB-like protein from *Paenibacillus apairius* strain DB482.

[00109] **SEQ ID NO:37** is an amino acid sequence from a TcaB-like protein from *Paenibacillus apairius* strain DB482.

[00110] **SEQ ID NO:38** is a nucleotide sequence from a tcaC-like gene from *Paenibacillus apairius* strain DB482.

[00111] **SEQ ID NO:39** is an amino acid sequence from a TcaC-like protein from *Paenibacillus apairius* strain DB482.

[00112] **SEQ ID NO:40** is a nucleotide sequence from a tccC-like gene from *Paenibacillus apairius* strain DB482.

[00113] **SEQ ID NO:41** is an amino acid sequence from a TccC-like protein from *Paenibacillus apairius* strain DB482.

[00114] **SEQ ID NO:42** is gene *tcdB1*, source organism *Photorhabdus* strain W14, gene designation *tcdB1-W14*.

[00115] **SEQ ID NO:43** is gene *tcdB2*, source organism *Photorhabdus* strain W14, gene designation *tcdB2-W14*.

[00116] **SEQ ID NO:44** is gene *tccC1*, source organism *Photorhabdus* strain W14, gene designation *tccC1-W14*.

[00117] **SEQ ID NO:45** is gene *tccC2*, source organism *Photorhabdus* strain W14, gene designation *tccC2-W14*.

[00118] **SEQ ID NO:46** is gene *tccC3*, source organism *Photorhabdus* strain W14, gene designation *tccC3-W14*.

[00119] **SEQ ID NO:47** is gene *tccC4*, source organism *Photorhabdus* strain W14, gene designation *tccC4-W14*.

[00120] **SEQ ID NO:48** is gene *tccC5*, source organism *Phtorhabdus* strain W14, gene designation *tccC5-W14*.

[00121] **SEQ ID NO:49** is the amino acid sequence of the XptA2 TC protein from *Xenorhabdus nematophilus* Xwi.

Detailed Description of the Invention

[00122] The subject invention provides unique biological alternatives for pest control. More specifically, the subject invention provides new sources of proteins that have toxin activity against insects, preferably lepidopterans, and other similar pests. The invention also relates to new sources of novel polynucleotides that can be used to encode such toxins, and to methods of making and methods of using the toxins and corresponding nucleic acid sequences to control insects and other like plant pests. The present invention addresses the need for novel insect control agents. The present invention relates to novel pesticidal proteins that are obtainable from *Paenibacillus,* and other, bacteria.

[00123] The subject invention initially stemmed from a discovery of a novel strain of *Paenibacillus*. This strain is referred to herein as DAS1529. To demonstrate the broad implications of this discovery, the discovery of another *Paenibacillus* strain is also exemplified. These strains have been deposited with the Agricultural Research Service Patent Culture Collection (NRRL) at 1815 North University Street Peoria, Ill. 61604 U.S.A. The deposited strains and the corresponding deposit dates and deposit numbers are as follows:

Deposited Strain	Deposit Date	Deposit Number
DAS1529	June 19, 2002	NRRL B-30599
DB482	June 17, 2003	NRRL B-30670

[00124] These cultures have been deposited for the purposes of this patent application and were deposited under conditions that assure that access to the cultures is available during the pendency of this patent application to one determined by the Commissioner of Patents and Trademarks to be entitled thereto under 37 CFR 1.14 and 35 U.S.C. 122. These deposits will be available as required by foreign patent laws in countries wherein counterparts of the subject application, or its progeny, are filed. However, it should be understood that the availability of a deposit does not constitute a license to practice the subject invention in derogation of patent rights granted by governmental action.

[00125] Further, the subject culture deposits were made in accordance with the provisions of the Budapest Treaty for the Deposit of Microorganisms, *i.e.*, they will be stored with all the care

necessary to keep them viable and uncontaminated for a period of at least five years after the most recent request for the furnishing of a sample of the deposit, and in any case, for a period of at least thirty (30) years after the date of deposit or for the enforceable life of any patent which may issue disclosing the culture. The depositor acknowledges the duty to replace the deposit should the depository be unable to furnish a sample when requested, due to the condition of the deposit. All restrictions on the availability to the public of the subject culture deposits will be irrevocably removed upon the granting of a patent disclosing them.

[00126] The discovery of the subject DAS1529 strain was surprising for a variety of reasons. This strain produces a unique, lepidopteran-toxic *Cry* protein. This strain, as well as DB482, also produce unique, toxin complex (TC)-like proteins (having some similarity to *Xenorhabdus/Photorhabdus* TCs). *Paenibacillus* isolate DB482 and toxins obtainable therefrom are highly preferred, and all are within the scope of the subject invention.

[00127] This is the first known disclosure of a *Paenibacillus* protein toxin having activity against a lepidopteran pest. The DAS1529 strain was found to have toxin activity against lepidopteran pests. This was a surprising discovery. Some species of *Paenibacillus* were known to have insecticidal activity against grubs/beetles/coleopterans. There have been no known reports of a *Paenibacillus* species or strain having toxicity to lepidopterans. Thus, the subject invention relates generally to *Paenibacillus* species that have activity against lepidopterans, and to screening *Paenibacillus* cultures, proteins therefrom, and libraries of clones therefrom, for activity against lepidopterans, and/or for genes that encode "lep toxins," and more particularly, for lepidopteran-toxic *Cry* proteins.

[00128] This is also the first known report of *Paenibacillus* having TC-like proteins. Thus, the subject invention relates to methods of screening *Paenibacillus* spp. for TC-like genes and proteins. It was very surprising to find that the DAS1529 and DB482 strains have TC-like operons and produce TC proteins (having some level of similarity to TC proteins of *Xenorhabdus* and *Photorhabdus*). TC proteins and genes identified herein were not heretofore known to exist in the genus *Paenibacillus*. This discovery broadens the scope of organisms (bacterial genera) in which TC protein genes have been found. Thus, the subject invention generally relates to TC proteins obtainable from *Paenibacillus* species, to methods of screening *Paenibacillus* species for such proteins, and the like. An example of a *Paenibacillus* species found using the methods of

the subject invention is *Paenibacillus apairius* strain DB482. This *P. apairius* strain also produces unique TC-like proteins.

[00129] While the subject TC proteins have some characteristics in common with TC proteins of *Xenorhabdus* and *Photorabdus*, the subject TC proteins are unique and different from previously known TC proteins. Thus, the subject application provides new classes of TC-like proteins and genes that encode these proteins obtainable from bacteria in the genera *Paenibacillus, Photorhabdus, Xenorhabdus, Serratia,* and the like.

[00130] The subject invention also relates to lepidopteran-toxic *Cry* proteins that are obtainable from *Paenibacillus* species. Thus, the subject invention relates to methods of screening *Paenibacillus* species for *cry* genes and *Cry* proteins that have toxin activity against a lepidopteran pest.

[00131] The DAS1529 *Cry* toxin is a very unique, *B.t.*-like *Cry* protein toxin. One other strain of *Paenibacillus*, a strain with activity against grubs, was known to produce a coleopteran-toxic *Cry* protein. That was a *Cry*18 protein, which was most related to *Cry*2 proteins (but only about 40% identity). The *Cry* protein exemplified herein shows only a low level of sequence identity and similarity to previously known *Cry* proteins. With that noted, of all the known *B.t. Cry* proteins, the subject *Cry* protein shares the most similarity to *Cry*1 proteins. One surprising aspect of the subject *Cry* protein is that it is very short, *i.e.*, even shorter than the *Cry*1Fa core toxin. The subject *Cry* protein has an identifiable Block 5 region at or near its C terminus. This toxin in its wild-type state has no protoxin portion, which is typically found on *Cry*1 toxins. The subject *Cry* toxin is surprisingly compressed. Thus, the subject invention generally relates to a new class of *Cry* proteins. This disclosure is also significant to the search for additional *cry* genes from *Bacillus thuringiensis* (*B.t.*). As would be clear to one skilled in the art having the benefit of the subject disclosure, other bacteria, such as *B.t.* and other *Bacillus* spp. (including *sphaericus*) could be screened for similar toxins and toxin genes. These methods of screening are within the scope of the subject invention.

[00132] The DAS1529 strain is the first known example of a natural bacterium that produces both a *Cry*-like toxin and TC-like proteins. Further surprising is that this is the first known example of a *cry* toxin gene being closely associated with (in genetic proximity to) TC protein genes. These pioneering observations thus enable one skilled in the art to screen appropriate species of bacteria for these types of unique operons and for these types of further components of known operons.

Such techniques are within the scope of the subject invention. The DAS1529 strain is an interesting example of a wild type strain having a TC-like operon with multiple TC protein genes of the same general type (*i.e.*, in this case, two tcaA-like and two tcaB-like genes). This could have implications for further gene discovery.

[00133] A further aspect of the subject invention stems from the surprising discovery that the *Paenibacillus* thiaminase protein has insecticidal activity. While this protein was known, it was in no way expected in the art that this enzyme would have exhibited toxin-like activity against insects/insect-like pests.

[00134] *Paenibacillus* TC proteins

[00135] More specifically regarding the exemplified TC proteins, the following TC proteins from strain DAS1529 have been fully characterized herein: two TcaA-like proteins ($TcaA_1$ and $TcaA_2$), two TcaB-like proteins ($TcaB_1$ and $TcaB_2$), a TcaC protein, and a TccC-like protein. The $TcaA_1$ and $TcaA_2$ proteins are highly similar to each other at the sequence level, and the $tcaB_1$ and $tcaB_2$ proteins are highly similar to each other at the sequence level. TC-like proteins obtainable from *Paenibacillus apairius* are also exemplified herein, and are within the scope of the subject invention.

[00136] The TC proteins of the subject invention can be used like other TC proteins. This would be readily apparent to one skilled in the art having the benefit of the subject disclosure when viewed in light of what was known in the art. *See, e.g.*, the Background section, above, which discusses R.H. ffrench-Constant and Bowen (2000) and U.S. Patent No. 6,048,838. For example, it was known that the Tca toxin complex of *Photorhabdus* is highly toxic to *Manduca sexta*.

[00137] While the exact molecular interactions of the TC proteins with each other, and their mechanism(s) of action, are not currently understood, some TC proteins were known to have "stand alone" insecticidal activity, and other TC proteins were known to enhance the activity of the stand-alone toxins produced by the same given organism. For example, it was known that the TcdA protein was active against *Manduca sexta*. TcaC and TccC, together, can be used to enhance the activity of TcdA. TcdB can be used (in place of TcaC) with TccC as a potentiator. TcbA is another *Photorhabdus* TC protein with stand-alone toxin activity. TcaC (or TcdB) together with TccC can also be used to enhance/potentiate the toxin activity of TcbA.

[00138] *Photorhabdus* TC proteins and "corresponding" TC proteins/genes from *Paenibacillus* are summarized below.

Photorhabdus	*Photorhabdus strain W14 nomenclature*	*Photorhabdus* Self homology	*Paenibacillus* 1529
TcaA	Toxin C	TccA	ORF3 (& 1)
TcaB		TccB	ORF4 (& 2)
TcaC		TcdB	ORF5
Tcb	Toxin B		
TccA	Toxin D	TcdA N terminus	
TccB		TcdA C terminus	
TccC			ORF6
TcdA	Toxin A	TccA + TccB	
TcdB		TcaC	

[00139] As indicated above, TccA has some level of homology with the N terminus of TcdA, and TccB has some level of homology with the C terminus of TcdA. Furthermore, TcdA is about 280 kDa, and TccA together with TccB are of about the same size, if combined, as TcdA. Furthermore, TcaA has some level of homology with TccA and likewise with the N terminus of TcdA. Still further, TcaB has some level of homology with TccB and likewise with the N terminus of TcdA. TccA and TcaA are of a similar size, as are TccB and TcaB.

[00140] Although some *Xenorhabdus* TC proteins were found to "correspond" to some of the *Photorhabdus* TC proteins, the "corresponding" proteins share only about 40% (approximately) sequence identity with each other. The subject TC proteins from *Paenibacillus* have about that same degree of sequence relatedness (~40% identity) with prior TC proteins.

[00141] As described in more detail below, one or more toxins of the subject invention can be used in combination with each other and/or with other toxins (*i.e.*, the *Photorhabdus* Tca complex was known to be active against *Manduca sexta*; various "combinations" of *Photorhabdus* TC proteins, for example, can be used together to enhance the activity of other, stand-alone *Photorhabdus* toxins; the use of *Photorhabdus* toxins "with" *B.t.* toxins, for example, has been proposed for resistance management.) Furthermore, *Paenibacillus* TC proteins of the subject invention are shown herein to be useful to enhance or potentiate the activity of a "stand-alone" *Xenorhabdus* toxin protein, for example. Provisional application No. 60/441,723 (Timothy D. Hey *et al.*), entitled "Mixing and Matching TC Proteins for Pest Control," relates to the surprising discovery that a TC protein derived from an organism of one genus such as *Photorhabdus* can be used interchangeably with a "corresponding" TC protein derived from an

organism of another genus. Further surprising data along these lines is presented below which further illustrate the utility of the *Paenibacillus* TC proteins of the subject invention. One reason that these results might be surprising is that there is only ~40% sequence identity between "corresponding" *Xhenorhabdus*, *Photorhabdus*, and the subject *Paenibacillus* TC proteins.

[00142] Proteins and toxins. The present invention provides easily administered, functional proteins. The present invention also provides a method for delivering insecticidal toxins that are functionally active and effective against many orders of insects, preferably lepidopteran insects. By "functional activity" (or "active against") it is meant herein that the protein toxins function as orally active insect control agents (alone or in combination with other proteins), that the proteins have a toxic effect (alone or in combination with other proteins), or are able to disrupt or deter insect growth and/or feeding which may or may not cause death of the insect. When an insect comes into contact with an effective amount of a "toxin" of the subject invention delivered via transgenic plant expression, formulated protein composition(s), sprayable protein composition(s), a bait matrix or other delivery system, the results are typically death of the insect, inhibition of the growth and/or proliferation of the insect, and/or prevention of the insects from feeding upon the source (preferably a transgenic plant) that makes the toxins available to the insects. Functional proteins of the subject invention can also enhance or improve the activity of other toxin proteins. Thus, terms such as "toxic," "toxicity," "toxin activity," and "pesticidally active" as used herein are meant to convey that the subject "toxins" have "functional activity" as defined herein.

[00143] Complete lethality to feeding insects is preferred, but is not required to achieve functional activity. If an insect avoids the toxin or ceases feeding, that avoidance will be useful in some applications, even if the effects are sublethal or lethality is delayed or indirect. For example, if insect resistant transgenic plants are desired, the reluctance of insects to feed on the plants is as useful as lethal toxicity to the insects because the ultimate objective is avoiding insect-induced plant damage.

[00144] There are many other ways in which toxins can be incorporated into an insect's diet. For example, it is possible to adulterate the larval food source with the toxic protein by spraying the food with a protein solution, as disclosed herein. Alternatively, the purified protein could be genetically engineered into an otherwise harmless bacterium, which could then be grown in culture, and either applied to the food source or allowed to reside in the soil in an area in which

insect eradication was desirable. Also, the protein could be genetically engineered directly into an insect food source. For instance, the major food source for many insect larvae is plant material. Therefore the genes encoding toxins can be transferred to plant material so that said plant material expresses the toxin of interest.

[00145] Transfer of the functional activity to plant or bacterial systems typically requires nucleic acid sequences, encoding the amino acid sequences for the toxins, integrated into a protein expression vector appropriate to the host in which the vector will reside. One way to obtain a nucleic acid sequence encoding a protein with functional activity is to isolate the native genetic material from the bacterial species which produce the toxins, using information deduced from the toxin's amino acid sequence, as disclosed herein. The native sequences can be optimized for expression in plants, for example, as discussed in more detail below. Optimized polynucleotide can also be designed based on the protein sequence.

[00146] The subject invention provides new classes of toxins having advantageous pesticidal activities. One way to characterize these classes of toxins and the polynucleotides that encode them is by defining a polynucleotide by its ability to hybridize, under a range of specified conditions, with an exemplified nucleotide sequence (the complement thereof and/or a probe or probes derived from either strand) and/or by their ability to be amplified by PCR using primers derived from the exemplified sequences.

[00147] There are a number of methods for obtaining the pesticidal toxins of the instant invention. For example, antibodies to the pesticidal toxins disclosed and claimed herein can be used to identify and isolate other toxins from a mixture of proteins. Specifically, antibodies may be raised to the portions of the toxins which are most constant and most distinct from other toxins. These antibodies can then be used to specifically identify equivalent toxins with the characteristic activity by immunoprecipitation, enzyme linked immunosorbent assay (ELISA), or western blotting. Antibodies to the toxins disclosed herein, or to equivalent toxins, or to fragments of these toxins, can be readily prepared using standard procedures. Monoclonal, polyclonal, specific, and/or cross-reactive antibodies can be made and used according to the subject invention. Such antibodies can be included in test kits for detecting the presence of proteins (and antigenic fragments thereof) of the subject invention.

[00148] One skilled in the art would readily recognize that toxins (and genes) of the subject invention can be obtained from a variety of sources. A toxin "from" or "obtainable from" the

subject DAS 1529 isolate and/or the *P. apiarius* isolate means that the toxin (or a similar toxin) can be obtained from this isolate or some other source, such as another bacterial strain or a transgenic plant. For example, one skilled in the art will readily recognize that, given the disclosure of a bacterial gene and toxin, a plant can be engineered to produce the toxin. Antibody preparations, nucleic acid probes (DNA and RNA), and the like may be prepared using the polynucleotide and/or amino acid sequences disclosed herein and used to screen and recover other toxin genes from other (natural) sources. Toxins of the subject invention can be obtained from a variety of sources/source microorganisms.

[00149] Polynucleotides and probes. The subject invention further provides nucleotide sequences that encode the toxins of the subject invention. The subject invention further provides methods of identifying and characterizing genes that encode pesticidal toxins. In one embodiment, the subject invention provides unique nucleotide sequences that are useful as hybridization probes and/or primers for PCR techniques. The primers produce characteristic gene fragments that can be used in the identification, characterization, and/or isolation of specific toxin genes. The nucleotide sequences of the subject invention encode toxins that are distinct from previously described toxins.

[00150] The polynucleotides of the subject invention can be used to form complete "genes" to encode proteins or peptides in a desired host cell. For example, as the skilled artisan would readily recognize, the subject polynucleotides can be appropriately placed under the control of a promoter in a host of interest, as is readily known in the art.

[00151] As the skilled artisan knows, DNA typically exists in a double-stranded form. In this arrangement, one strand is complementary to the other strand and vice versa. As DNA is replicated in a plant (for example), additional complementary strands of DNA are produced. The "coding strand" is often used in the art to refer to the strand that binds with the anti-sense strand. The mRNA is transcribed from the "anti-sense" strand of DNA. The "sense" or "coding" strand has a series of codons (a codon is three nucleotides that can be read as a three-residue unit to specify a particular amino acid) that can be read as an open reading frame (ORF) to form a protein or peptide of interest. In order to express a protein *in vivo*, a strand of DNA is typically transcribed into a complementary strand of mRNA which is used as the template for the protein. Thus, the subject invention includes the use of the exemplified polynucleotides shown in the attached sequence listing and/or equivalents including the complementary strands. RNA and

PNA (peptide nucleic acids) that are functionally equivalent to the exemplified DNA are included in the subject invention.

[00152] In one embodiment of the subject invention, bacterial isolates can be cultivated under conditions resulting in high multiplication of the microbe. After treating the microbe to provide single-stranded genomic nucleic acid, the DNA can be contacted with the primers of the invention and subjected to PCR amplification. Characteristic fragments of toxin-encoding genes will be amplified by the procedure, thus identifying the presence of the toxin-encoding gene(s).

[00153] Further aspects of the subject invention include genes and isolates identified using the methods and nucleotide sequences disclosed herein. The genes thus identified encode toxins active against pests.

[00154] Toxins and genes of the subject invention can be identified and obtained by using oligonucleotide probes, for example. These probes are detectable nucleotide sequences which may be detectable by virtue of an appropriate label or may be made inherently fluorescent as described in International Application No. WO 93/16094. The probes (and the polynucleotides of the subject invention) may be DNA, RNA, or PNA. In addition to adenine (A), cytosine (C), guanine (G), thymine (T), and uracil (U; for RNA molecules), synthetic probes (and polynucleotides) of the subject invention can also have inosine (a neutral base capable of pairing with all four bases; sometimes used in place of a mixture of all four bases in synthetic probes). Thus, where a synthetic, degenerate oligonucleotide is referred to herein, and "n" is used generically, "n" can be G, A, T, C, or inosine. Ambiguity codes as used herein are in accordance with standard IUPAC naming conventions as of the filing of the subject application (for example, R means A or G, Y means C or T, etc.).

[00155] As is well known in the art, if a probe molecule hybridizes with a nucleic acid sample , it can be reasonably assumed that the probe and sample have substantial homology/similarity/identity. Preferably, hybridization of the polynucleotide is first conducted followed by washes under conditions of low, moderate, or high stringency by techniques well-known in the art, as described in, for example, Keller, G.H., M.M. Manak (1987) *DNA Probes*, Stockton Press, New York, NY, pp. 169-170. For example, as stated therein, low stringency conditions can be achieved by first washing with 2x SSC (Standard Saline Citrate)/0.1% SDS (Sodium Dodecyl Sulfate) for 15 minutes at room temperature. Two washes are typically performed. Higher stringency can then be achieved by lowering the salt concentration and/or by

raising the temperature. For example, the wash described above can be followed by two washings with 0.1x SSC/0.1% SDS for 15 minutes each at room temperature followed by subsequent washes with 0.1x SSC/0.1% SDS for 30 minutes each at 55° C. These temperatures can be used with other hybridization and wash protocols set forth herein and as would be known to one skilled in the art (SSPE can be used as the salt instead of SSC, for example). The 2x SSC/0.1% SDS can be prepared by adding 50 ml of 20x SSC and 5 ml of 10% SDS to 445 ml of water. 20x SSC can be prepared by combining NaCl (175.3 g/0.150 M), sodium citrate (88.2 g/0.015 M), and water to 1 liter, followed by adjusting pH to 7.0 with 10 N NaOH. 10% SDS can be prepared by dissolving 10 g of SDS in 50 ml of autoclaved water, diluting to 100 ml, and aliquotting.

[00156] Detection of the probe provides a means for determining in a known manner whether hybridization has been maintained. Such a probe analysis provides a rapid method for identifying toxin-encoding genes of the subject invention. The nucleotide segments which are used as probes according to the invention can be synthesized using a DNA synthesizer and standard procedures. These nucleotide sequences can also be used as PCR primers to amplify genes of the subject invention.

[00157] Probes for use according to the subject invention can be derived from a variety of sources, such as any gene mentioned or suggested herein. For example, all or part of any of the following types of genes (coding and/or noncoding or complementary strands thereof) can be used according to the subject invention: tcaA, tcaB, tcaC, tcbA, tccA, tccB, tccC, tcdA, tcdB, xptA1, xptD1, xptB1, xptC1, xptA2, sepA, sepB, and sepC. Unless specifically indicated otherwise, reference to a "tccC" gene, for example, includes all specific alleles (such as tccC1 and tccC2) of this type of gene. The same is true for all the other genes (*e.g.*, tcdB2, tccC3, and the alleles mentioned in Table 17).

[00158] Hybridization characteristics of a molecule can be used to define polynucleotides of the subject invention. Thus the subject invention includes polynucleotides (and/or their complements, preferably their full complements) that hybridize with a polynucleotide (or an oligonucleotide or primer) exemplified or suggested herein.

[00159] As used herein "stringent" conditions for hybridization refers to conditions which achieve the same, or about the same, degree of specificity of hybridization as the conditions employed by the current applicants. Specifically, hybridization of immobilized DNA on Southern blots with

^{32}P-labeled gene-specific probes was performed by standard methods (*see, e.g.*, Maniatis, T., E.F. Fritsch, J. Sambrook [1982] *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY). In general, hybridization and subsequent washes were carried out under conditions that allowed for detection of target sequences. For double-stranded DNA gene probes, hybridization was carried out overnight at 20-25° C below the melting temperature (Tm) of the DNA hybrid in 6x SSPE, 5x Denhardt's solution, 0.1% SDS, 0.1 mg/ml denatured DNA. The melting temperature is described by the following formula (Beltz, G.A., K.A. Jacobs, T.H. Eickbush, P.T. Cherbas, and F.C. Kafatos [1983] *Methods of Enzymology*, R. Wu, L. Grossman and K. Moldave [eds.] Academic Press, New York 100:266-285):

Tm = 81.5° C + 16.6 Log[Na+] + 0.41(%G+C) - 0.61(%formamide) - 600/length of duplex in base pairs.

[00160] Washes are typically carried out as follows:

(1) Twice at room temperature for 15 minutes in 1x SSPE, 0.1% SDS (low stringency wash).

(2) Once at Tm-20°C for 15 minutes in 0.2x SSPE, 0.1% SDS (moderate stringency wash).

[00161] For oligonucleotide probes, hybridization was carried out overnight at 10-20°C below the melting temperature (Tm) of the hybrid in 6x SSPE, 5x Denhardt's solution, 0.1% SDS, 0.1 mg/ml denatured DNA. Tm for oligonucleotide probes was determined by the following formula:

Tm (° C) = 2(number T/A base pairs) + 4(number G/C base pairs)

(Suggs, S.V., T. Miyake, E.H. Kawashime, M.J. Johnson, K. Itakura, and R.B. Wallace [1981] *ICN-UCLA Symp. Dev. Biol. Using Purified Genes*, D.D. Brown [ed.], Academic Press, New York, 23:683-693).

[00162] Washes were typically carried out as follows:

(1) Twice at room temperature for 15 minutes 1x SSPE, 0.1% SDS (low stringency wash).

(2) Once at the hybridization temperature for 15 minutes in 1x SSPE, 0.1% SDS (moderate stringency wash).

[00163] In general, salt and/or temperature can be altered to change stringency. With a labeled DNA fragment >70 or so bases in length, the following conditions can be used:

Low:	1 or 2x SSPE, room temperature
Low:	1 or 2x SSPE, 42° C
Moderate:	0.2x or 1x SSPE, 65° C
High:	0.1x SSPE, 65° C.

[00164] Duplex formation and stability depend on substantial complementarity between the two strands of a hybrid, and, as noted above, a certain degree of mismatch can be tolerated. Therefore, the probe sequences of the subject invention include mutations (both single and multiple), deletions, insertions of the described sequences, and combinations thereof, wherein said mutations, insertions and deletions permit formation of stable hybrids with the target polynucleotide of interest. Mutations, insertions, and deletions can be produced in a given polynucleotide sequence in many ways, and these methods are known to an ordinarily skilled artisan. Other methods may become known in the future.

[00165] PCR technology. Polymerase Chain Reaction (PCR) is a repetitive, enzymatic, primed synthesis of a nucleic acid sequence. This procedure is well-known and commonly used by those skilled in this art (*see* Mullis, U.S. Patent Nos. 4,683,195, 4,683,202, and 4,800,159; Saiki, Randall K., Stephen Scharf, Fred Faloona, Kary B. Mullis, Glenn T. Horn, Henry A. Erlich, Norman Arnheim [1985] "Enzymatic Amplification of β-Globin Genomic Sequences and Restriction Site Analysis for Diagnosis of Sickle Cell Anemia," *Science* 230:1350-1354). PCR is based on the enzymatic amplification of a DNA fragment of interest that is flanked by two oligonucleotide primers that hybridize to opposite strands of the target sequence. The primers are oriented with the 3′ ends pointing towards each other. Repeated cycles of heat denaturation of the template, annealing of the primers to their complementary sequences, and extension of the annealed primers with a DNA polymerase result in the amplification of the segment defined by the 5′ ends of the PCR primers. The extension product of each primer can serve as a template for the other primer, so each cycle essentially doubles the amount of DNA fragment produced in the previous cycle. This results in the exponential accumulation of the specific target fragment, up to several million-fold in a few hours. By using a thermostable DNA polymerase such as *Taq* polymerase, isolated from the thermophilic bacterium *Thermus aquaticus*, the amplification process can be completely automated. Other enzymes which can be used are known to those skilled in the art.

[00166] The DNA sequences of the subject invention can be used as primers for PCR

amplification. In performing PCR amplification, a certain degree of mismatch can be tolerated between primer and template. Therefore, mutations, deletions, and insertions (especially additions of nucleotides to the 5′ end) of the exemplified primers fall within the scope of the subject invention. Mutations, insertions, and deletions can be produced in a given primer by methods known to an ordinarily skilled artisan.

[00167] Modification of genes and toxins. The genes and toxins useful according to the subject invention include not only the specifically exemplified full-length sequences, but also portions, segments and/or fragments (including internal and/or terminal deletions compared to the full-length molecules) of these sequences, variants, mutants, chimerics, and fusions thereof. For example, toxins of the subject invention may be used in the form of chimeric toxins produced by combining portions of two or more toxins/proteins.

[00168] Proteins of the subject invention can have substituted amino acids so long as they retain the characteristic pesticidal/ functional activity of the proteins specifically exemplified herein. "Variant" genes have nucleotide sequences that encode the same toxins or equivalent toxins having pesticidal activity equivalent to an exemplified protein. The terms "variant proteins" and "equivalent toxins" refer to toxins having the same or essentially the same biological/functional activity against the target pests and equivalent sequences as the exemplified toxins. As used herein, reference to an "equivalent" sequence refers to sequences having amino acid substitutions, deletions, additions, or insertions which improve or do not adversely affect pesticidal activity. Fragments retaining pesticidal activity are also included in this definition. Fragments and other equivalents that retain the same or similar function, or "toxin activity," of a corresponding fragment of an exemplified toxin are within the scope of the subject invention. Changes, such as amino acid substitutions or additions, can be made for a variety of purposes, such as increasing (or decreasing) protease stability of the protein (without materially/substantially decreasing the functional activity of the toxin).

[00169] Equivalent toxins and/or genes encoding these equivalent toxins can be obtained/derived from wild-type or recombinant bacteria and/or from other wild-type or recombinant organisms using the teachings provided herein. Other *Bacillus*, *Paenibacillus*, *Photorhabdus*, and *Xenorhabdus* species, for example, can be used as source isolates.

[00170] Variations of genes may be readily constructed using standard techniques for making point mutations, for example. In addition, U.S. Patent No. 5,605,793, for example, describes

methods for generating additional molecular diversity by using DNA reassembly after random fragmentation. Variant genes can be used to produce variant proteins; recombinant hosts can be used to produce the variant proteins. Using these "gene shuffling" techniques, equivalent genes and proteins can be constructed that comprise any 5, 10, or 20 contiguous residues (amino acid or nucleotide) of any sequence exemplified herein. As one skilled in the art knows, the gene shuffling techniques can be adjusted to obtain equivalents having, for example, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, or 500 or more contiguous residues

(amino acid or nucleotide), corresponding to a segment (of the same size) in any of the exemplified sequences (or the complements (full complements) thereof). Similarly sized segments, especially those for conserved regions, can also be used as probes and/or primers.

[00171] Fragments of full-length genes can be made using commercially available exonucleases or endonucleases according to standard procedures. For example, enzymes such as *Bal*31 or site-directed mutagenesis can be used to systematically cut off nucleotides from the ends of these genes. Also, genes which encode active fragments may be obtained using a variety of restriction enzymes. Proteases may be used to directly obtain active fragments of these toxins.

[00172] It is within the scope of the invention as disclosed herein that toxins may be truncated and still retain functional activity. By "truncated toxin" is meant that a portion of a toxin protein may be cleaved and yet still exhibit activity after cleavage. Cleavage can be achieved by proteases inside or outside of the insect gut. Furthermore, effectively cleaved proteins can be produced using molecular biology techniques wherein the DNA bases encoding said toxin are removed either through digestion with restriction endonucleases or other techniques available to the skilled artisan. After truncation, said proteins can be expressed in heterologous systems such as *E. coli*, baculoviruses, plant-based viral systems, yeast and the like and then placed in insect assays as disclosed herein to determine activity. It is well-known in the art that truncated toxins can be successfully produced so that they retain functional activity while having less than the entire, full-length sequence. It is well known in the art that B.t. toxins can be used in a truncated (core toxin) form. *See, e.g.*, Adang *et al.*, *Gene* 36:289-300 (1985), "Characterized full-length and truncated plasmid clones of the crystal protein of *Bacillus thuringiensis* subsp *kurstaki* HD-73 and their toxicity to *Manduca sexta.*" There are other examples of truncated proteins that retain insecticidal activity, including the insect juvenile hormone esterase (U.S. Pat. No. 5,674,485 to the Regents of the University of California). As used herein, the term "toxin" is also meant to include functionally active truncations. On the other hand, a protoxin portion (typically the C-terminal half of a typical *B.t. Cry* toxin) can be added to form an active, full-length protein. *See, e.g.,* U.S. Patent No. 6,218,188.

[00173] Certain toxins of the subject invention have been specifically exemplified herein. As these toxins are merely exemplary of the toxins of the subject invention, it should be readily apparent that the subject invention comprises variant or equivalent toxins (and nucleotide sequences coding for equivalent toxins) having the same or similar pesticidal activity of the

exemplified toxin. Equivalent toxins will have amino acid similarity (and/or homology) with an exemplified toxin. The amino acid identity will typically be greater than 60%, preferably greater than 75%, more preferably greater than 80%, even more preferably greater than 90%, and can be greater than 95%. Preferred polynucleotides and proteins of the subject invention can also be defined in terms of more particular identity and/or similarity ranges. For example, the identity and/or similarity can be 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% as compared to a sequence exemplified herein. Unless otherwise specified, as used herein percent sequence identity and/or similarity of two nucleic acids is determined using the algorithm of Karlin and Altschul (1990), *Proc. Natl. Acad. Sci. USA* 87:2264-2268, modified as in Karlin and Altschul (1993), *Proc. Natl. Acad. Sci. USA* 90:5873-5877. Such an algorithm is incorporated into the NBLAST and XBLAST programs of Altschul *et al.* (1990), *J. Mol. Biol.* 215:402-410. BLAST nucleotide searches are performed with the NBLAST program, score = 100, wordlength = 12. To obtain gapped alignments for comparison purposes, Gapped BLAST is used as described in Altschul *et al.* (1997), *Nucl. Acids Res.* 25:3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (NBLAST and XBLAST) are used. *See* NCBI/NIH website. The scores can also be calculated using the methods and algorithms of Crickmore *et al.* as described in the Background section, above.

[00174] The amino acid homology/similarity/identity will be highest in critical regions of the toxin which account for biological activity or are involved in the determination of three-dimensional configuration which is ultimately responsible for the biological activity. In this regard, certain amino acid substitutions are acceptable and can be expected to be tolerated. For example, these substitutions can be in regions of the protein that are not critical to activity. Analyzing the crystal structure of a protein, and software-based protein structure modeling, can be used to identify regions of a protein that can be modified (using site-directed mutagenesis, shuffling, etc.) to actually change the properties and/or increase the functionality of the protein.

[00175] Various properties and targeted 3D features of the protein can also be changed without adversely affecting the toxin activity/functionality of the protein. Conservative amino acid substitutions can be expected to be tolerated/to not adversely affect the three-dimensional configuration of the molecule. Amino acids can be placed in the following classes: non-polar,

uncharged polar, basic, and acidic. Conservative substitutions whereby an amino acid of one class is replaced with another amino acid of the same type fall within the scope of the subject invention so long as the substitution is not adverse to the biological activity of the compound. **Table 1** provides a listing of examples of amino acids belonging to each class.

Table 1.

Class of Amino Acid	Examples of Amino Acids
Nonpolar	Ala, Val, Leu, Ile, Pro, Met, Phe, Trp
Uncharged Polar	Gly, Ser, Thr, Cys, Tyr, Asn, Gln
Acidic	Asp, Glu
Basic	Lys, Arg, His

[00176] In some instances, non-conservative substitutions can also be made. The critical factor is that these substitutions must not significantly detract from the functional/biological activity of the toxin.

[00177] As used herein, reference to "isolated" polynucleotides and/or "purified" toxins refers to these molecules when they are not associated with the other molecules with which they would be found in nature. Thus, reference to "isolated" and/or "purified" signifies the involvement of the "hand of man" as described herein. For example, a bacterial toxin "gene" of the subject invention put into a plant for expression is an "isolated polynucleotide." Likewise, a *Paenibacillus* protein, exemplified herein, produced by a plant is an "isolated protein."

[00178] Because of the degeneracy/redundancy of the genetic code, a variety of different DNA sequences can encode the amino acid sequences disclosed herein. It is well within the skill of a person trained in the art to create alternative DNA sequences that encode the same, or essentially the same, toxins. These variant DNA sequences are within the scope of the subject invention.

[00179] Optimization of sequence for expression in plants. To obtain high expression of heterologous genes in plants it may be preferred to reengineer said genes so that they are more efficiently expressed in (the cytoplasm of) plant cells. Maize is one such plant where it may be preferred to re-design the heterologous gene(s) prior to transformation to increase the expression level thereof in said plant. Therefore, an additional step in the design of genes encoding a bacterial toxin is reengineering of a heterologous gene for optimal expression.

[00180] One reason for the reengineering of a bacterial toxin for expression in maize is due to the non-optimal G+C content of the native gene. For example, the very low G+C content of many native bacterial gene(s) (and consequent skewing towards high A+T content) results in the generation of sequences mimicking or duplicating plant gene control sequences that are known to be highly A+T rich. The presence of some A+T-rich sequences within the DNA of gene(s) introduced into plants (*e.g.*, TATA box regions normally found in gene promoters) may result in aberrant transcription of the gene(s). On the other hand, the presence of other regulatory sequences residing in the transcribed mRNA (*e.g.*, polyadenylation signal sequences (AAUAAA), or sequences complementary to small nuclear RNAs involved in pre-mRNA splicing) may lead to RNA instability. Therefore, one goal in the design of genes encoding a bacterial toxin for maize expression, more preferably referred to as plant optimized gene(s), is to generate a DNA sequence having a higher G+C content, and preferably one close to that of maize genes coding for metabolic enzymes. Another goal in the design of the plant optimized gene(s) encoding a bacterial toxin is to generate a DNA sequence in which the sequence modifications do not hinder translation.

[00181] The table below (**Table 2**) illustrates how high the G+C content is in maize. For the data in **Table 2**, coding regions of the genes were extracted from GenBank (Release 71) entries, and base compositions were calculated using the MacVector™ program (Accelerys, Burlington, MA). Intron sequences were ignored in the calculations.

[00182] Due to the plasticity afforded by the redundancy/degeneracy of the genetic code (i.e., some amino acids are specified by more than one codon), evolution of the genomes in different organisms or classes of organisms has resulted in differential usage of redundant codons. This "codon bias" is reflected in the mean base composition of protein coding regions. For example, organisms with relatively low G+C contents utilize codons having A or T in the third position of redundant codons, whereas those having higher G+C contents utilize codons having G or C in the third position. It is thought that the presence of "minor" codons within a mRNA may reduce the absolute translation rate of that mRNA, especially when the relative abundance of the charged tRNA corresponding to the minor codon is low. An extension of this is that the diminution of translation rate by individual minor codons would be at least additive for multiple minor codons. Therefore, mRNAs having high relative contents of minor codons would have

correspondingly low translation rates. This rate would be reflected by subsequent low levels of the encoded protein.

[00183] In reengineering genes encoding a bacterial toxin for maize (or other plant, such as cotton or soybean) expression, the codon bias of the plant has been determined. The codon bias for maize is the statistical codon distribution that the plant uses for coding its proteins and the preferred codon usage is shown in **Table 3.** After determining the bias, the percent frequency of the codons in the gene(s) of interest is determined. The primary codons preferred by the plant should be determined as well as the second and third choice of preferred codons. Afterwards, the amino acid sequence of the bacterial toxin of interest is reverse translated so that the resulting nucleic acid sequence codes for exactly the same protein as the native gene wanting to be heterologously expressed. The new DNA sequence is designed using codon bias information so that it corresponds to the most preferred codons of the desired plant. The new sequence is then analyzed for restriction enzyme sites that might have been created by the modification. The identified sites are further modified by replacing the codons with second or third choice preferred codons. Other sites in the sequence which could affect transcription or translation of the gene of interest are the exon:intron junctions (5' or 3'), poly A addition signals, or RNA polymerase termination signals. The sequence is further analyzed and modified to reduce the frequency of TA or GC doublets. In addition to the doublets, G or C sequence blocks that have more than about four residues that are the same can affect transcription of the sequence. Therefore, these blocks are also modified by replacing the codons of first or second choice, etc. with the next preferred codon of choice.

Table 2

Compilation of G + C contents of protein coding regions of maize genes

Protein Class.sup.a	Range % G + C	Mean % G + C.sup.b
Metabolic Enzymes (76)	44.4-75.3	59.0 (.+-.8.0)
Structural Proteins (18)	48.6-70.5	63.6 (.+-.6.7)
Regulatory Proteins (5)	57.2-68.8	62.0 (.+-.4.9)
Uncharacterized Proteins (9)	41.5-70.3	64.3 (.+-.7.2)
All Proteins (108)	44.4-75.3	60.8 (.+-.5.2)

.sup.a Number of genes in class given in parentheses.
.sup.b Standard deviations given in parentheses.
.sup.c Combined groups mean ignored in mean calculation

[00184] It is preferred that the plant optimized gene(s) encoding a bacterial toxin contain about

63% of first choice codons, between about 22% to about 37% second choice codons, and between about 15% to about 0% third choice codons, wherein the total percentage is 100%. Most preferred the plant optimized gene(s) contains about 63% of first choice codons, at least about 22% second choice codons, about 7.5% third choice codons, and about 7.5% fourth choice codons, wherein the total percentage is 100%. The preferred codon usage for engineering genes for maize expression are shown in **Table 3**. The method described above enables one skilled in the art to modify gene(s) that are foreign to a particular plant so that the genes are optimally expressed in plants. The method is further illustrated in PCT application WO 97/13402.

[00185] In order to design plant optimized genes encoding a bacterial toxin, the amino acid sequence of said protein is reverse translated into a DNA sequence utilizing a non-redundant genetic code established from a codon bias table compiled for the gene sequences for the particular plant, as shown in **Table 2.** The resulting DNA sequence, which is completely homogeneous in codon usage, is further modified to establish a DNA sequence that, besides having a higher degree of codon diversity, also contains strategically placed restriction enzyme recognition sites, desirable base composition, and a lack of sequences that might interfere with transcription of the gene, or translation of the product mRNA.

Table 3.

Preferred amino acid codons for proteins expressed in maize

Amino Acid	Codon*
Alanine	GCC/GCG
Cysteine	TGC/TGT
Aspartic Acid	GAC/GAT
Glutamic Acid	GAG/GAA
Phenylalanine	TTC/TTT
Glycine	GGC/GGG
Histidine	CAC/CAT
Isoleucine	ATC/ATT
Lysine	AAG/AAA
Leucine	CTG/CTC
Methionine	ATG
Asparagine	AAC/AAT
Proline	CCG/CCA
Glutamine	CAG/CAA
Arginine	AGG/CGC
Serine	AGC/TCC
Threonine	ACC/ACG
Valine	GTG/GTC
Tryptophan	TGG
Tryrosine	TAC/TAT
Stop	TGA/TAG

*The first and second preferred codons for maize.

[00186] Thus, synthetic genes that are functionally equivalent to the toxins/genes of the subject invention can be used to transform hosts, including plants. Additional guidance regarding the production of synthetic genes can be found in, for example, U.S. Patent No. 5,380,831.

[00187] In some cases, especially for expression in plants, it can be advantageous to use truncated genes that express truncated proteins. Höfte *et al.* 1989, for example, discussed in the Background Section above, discussed protoxin and core toxin segments of *B.t.* toxins. Preferred truncated genes will typically encode 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% of the full-length toxin.

[00188] Transgenic hosts. The toxin-encoding genes of the subject invention can be introduced into a wide variety of microbial or plant hosts. In preferred embodiments, transgenic plant cells and plants are used. Preferred plants (and plant cells) are corn, maize, and cotton.

[00189] In preferred embodiments, expression of the toxin gene results, directly or indirectly, in

the intracellular production (and maintenance) of the pesticide proteins. Plants can be rendered insect-resistant in this manner. When transgenic/recombinant/transformed/transfected host cells (or contents thereof) are ingested by the pests, the pests will ingest the toxin. This is the preferred manner in which to cause contact of the pest with the toxin. The result is control (killing or making sick) of the pest. Sucking pests can also be controlled in a similar manner. Alternatively, suitable microbial hosts, *e.g.*, *Pseudomonas* such as *P. fluorescens*, can be applied where target pests are present; the microbes can proliferate there, and are ingested by the target pests. The microbe hosting the toxin gene can be treated under conditions that prolong the activity of the toxin and stabilize the cell. The treated cell, which retains the toxic activity, can then be applied to the environment of the target pest.

[00190] Where the toxin gene is introduced via a suitable vector into a microbial host, and said host is applied to the environment in a living state, certain host microbes should be used. Microorganism hosts are selected which are known to occupy the "phytosphere" (phylloplane, phyllosphere, rhizosphere, and/or rhizoplane) of one or more crops of interest. These microorganisms are selected so as to be capable of successfully competing in the particular environment (crop and other insect habitats) with the wild-type microorganisms, provide for stable maintenance and expression of the gene expressing the polypeptide pesticide, and, desirably, provide for improved protection of the pesticide from environmental degradation and inactivation.

[00191] A large number of microorganisms are known to inhabit the phylloplane (the surface of the plant leaves) and/or the rhizosphere (the soil surrounding plant roots) of a wide variety of important crops. These microorganisms include bacteria, algae, and fungi. Of particular interest are microorganisms, such as bacteria, *e.g.*, genera *Pseudomonas*, *Erwinia*, *Serratia*, *Klebsiella*, *Xanthomonas*, *Streptomyces*, *Rhizobium*, *Rhodopseudomonas*, *Methylophilius*, *Agrobacterium*, *Acetobacter*, *Lactobacillus*, *Arthrobacter*, *Azotobacter*, *Leuconostoc*, and *Alcaligenes*; fungi, particularly yeast, *e.g.*, genera *Saccharomyces*, *Cryptococcus*, *Kluyveromyces*, *Sporobolomyces*, *Rhodotorula*, and *Aureobasidium*. Of particular interest are such phytosphere bacterial species as *Pseudomonas syringae*, *Pseudomonas fluorescens*, *Serratia marcescens*, *Acetobacter xylinum*, *Agrobacterium tumefaciens*, *Rhodopseudomonas spheroides*, *Xanthomonas campestris*, *Rhizobium melioti*, *Alcaligenes entrophus*, and *Azotobacter vinlandii*; and phytosphere yeast species such as *Rhodotorula rubra*, *R. glutinis*, *R. marina*, *R. aurantiaca*, *Cryptococcus albidus*,

C. diffluens, *C. laurentii*, *Saccharomyces rosei*, *S. pretoriensis*, *S. cerevisiae*, *Sporobolomyces roseus*, *S. odorus*, *Kluyveromyces veronae*, and *Aureobasidium pollulans*. Also of interest are pigmented microorganisms.

[00192] Insertion of genes to form transgenic hosts. One aspect of the subject invention is the transformation/transfection of plants, plant cells, and other host cells with polynucleotides of the subject invention that express proteins of the subject invention. Plants transformed in this manner can be rendered resistant to attack by the target pest(s).

[00193] A wide variety of methods are available for introducing a gene encoding a pesticidal protein into the target host under conditions that allow for stable maintenance and expression of the gene. These methods are well known to those skilled in the art and are described, for example, in United States Patent No. 5,135,867.

[00194] For example, a large number of cloning vectors comprising a replication system in *E. coli* and a marker that permits selection of the transformed cells are available for preparation for the insertion of foreign genes into higher plants. The vectors comprise, for example, pBR322, pUC series, M13mp series, pACYC184, etc. Accordingly, the sequence encoding the toxin can be inserted into the vector at a suitable restriction site. The resulting plasmid is used for transformation into *E. coli*. The *E. coli* cells are cultivated in a suitable nutrient medium, then harvested and lysed. The plasmid is recovered. Sequence analysis, restriction analysis, electrophoresis, and other biochemical-molecular biological methods are generally carried out as methods of analysis. After each manipulation, the DNA sequence used can be cleaved and joined to the next DNA sequence. Each plasmid sequence can be cloned in the same or other plasmids. Depending on the method of inserting desired genes into the plant, other DNA sequences may be necessary. If, for example, the Ti or Ri plasmid is used for the transformation of the plant cell, then at least the right border, but often the right and the left border of the Ti or Ri plasmid T-DNA, has to be joined as the flanking region of the genes to be inserted. The use of T-DNA for the transformation of plant cells has been intensively researched and described in EP 120 516; Hoekema (1985) In: *The Binary Plant Vector System*, Offset-durkkerij Kanters B.V., Alblasserdam, Chapter 5; Fraley *et al.*, *Crit. Rev. Plant Sci.* 4:1-46; and An *et al.* (1985) *EMBO J.* 4:277-287.

[00195] A large number of techniques are available for inserting DNA into a plant host cell. Those techniques include transformation with T-DNA using *Agrobacterium tumefaciens* or

Agrobacterium rhizogenes as transformation agent, fusion, injection, biolistics (microparticle bombardment), or electroporation as well as other possible methods. If *Agrobacteria* are used for the transformation, the DNA to be inserted has to be cloned into special plasmids, namely either into an intermediate vector or into a binary vector. The intermediate vectors can be integrated into the Ti or Ri plasmid by homologous recombination owing to sequences that are homologous to sequences in the T-DNA. The Ti or Ri plasmid also comprises the vir region necessary for the transfer of the T-DNA. Intermediate vectors cannot replicate themselves in *Agrobacteria*. The intermediate vector can be transferred into *Agrobacterium tumefaciens* by means of a helper plasmid (conjugation). Binary vectors can replicate themselves both in *E. coli* and in *Agrobacteria*. They comprise a selection marker gene and a linker or polylinker which are framed by the right and left T-DNA border regions. They can be transformed directly into *Agrobacteria* (Holsters *et al.* [1978] *Mol. Gen. Genet.* 163:181-187). The *Agrobacterium* used as host cell is to comprise a plasmid carrying a vir region. The vir region is necessary for the transfer of the T-DNA into the plant cell. Additional T-DNA may be contained. The bacterium so transformed is used for the transformation of plant cells. Plant explants can advantageously be cultivated with *Agrobacterium tumefaciens* or *Agrobacterium rhizogenes* for the transfer of the DNA into the plant cell. Whole plants can then be regenerated from the infected plant material (for example, pieces of leaf, segments of stalk, roots, but also protoplasts or suspension-cultivated cells) in a suitable medium, which may contain antibiotics or biocides for selection. The plants so obtained can then be tested for the presence of the inserted DNA. No special demands are made of the plasmids in the case of injection and electroporation. It is possible to use ordinary plasmids, such as, for example, pUC derivatives.

[00196] The transformed cells grow inside the plants in the usual manner. They can form germ cells and transmit the transformed trait(s) to progeny plants. Such plants can be grown in the normal manner and crossed with plants that have the same transformed hereditary factors or other hereditary factors. The resulting hybrid individuals have the corresponding phenotypic properties.

[00197] In some preferred embodiments of the invention, genes encoding the bacterial toxin are expressed from transcriptional units inserted into the plant genome. Preferably, said transcriptional units are recombinant vectors capable of stable integration into the plant genome and enable selection of transformed plant lines expressing mRNA encoding the proteins.

[00198] Once the inserted DNA has been integrated in the genome, it is relatively stable there (and does not come out again). It normally contains a selection marker that confers on the transformed plant cells resistance to a biocide or an antibiotic, such as kanamycin, G418, bleomycin, hygromycin, or chloramphenicol, *inter alia*. The individually employed marker should accordingly permit the selection of transformed cells rather than cells that do not contain the inserted DNA. The gene(s) of interest are preferably expressed either by constitutive or inducible promoters in the plant cell. Once expressed, the mRNA is translated into proteins, thereby incorporating amino acids of interest into protein. The genes encoding a toxin expressed in the plant cells can be under the control of a constitutive promoter, a tissue-specific promoter, or an inducible promoter.

[00199] Several techniques exist for introducing foreign recombinant vectors into plant cells, and for obtaining plants that stably maintain and express the introduced gene. Such techniques include the introduction of genetic material coated onto microparticles directly into cells (U.S. Pat. Nos. 4,945,050 to Cornell and 5,141,131 to DowElanco, now Dow AgroSciences, LLC). In addition, plants may be transformed using *Agrobacterium* technology, *see* U.S. Pat. No. 5,177,010 to University of Toledo; 5,104,310 to Texas A&M; European Patent Application 0131624B1; European Patent Applications 120516, 159418B1 and 176,112 to Schilperoot; U.S. Pat. Nos. 5,149,645, 5,469,976, 5,464,763 and 4,940,838 and 4,693,976 to Schilperoot; European Patent Applications 116718, 290799, 320500 all to Max Planck; European Patent Applications 604662 and 627752, and U.S. Pat. No. 5,591,616, to Japan Tobacco; European Patent Applications 0267159 and 0292435, and U.S. Pat. No. 5,231,019, all to Ciba Geigy, now Novartis; U.S. Pat. Nos. 5,463,174 and 4,762,785, both to Calgene; and U.S. Pat. Nos. 5,004,863 and 5,159,135, both to Agracetus. Other transformation technology includes whiskers technology. *See* U.S. Pat. Nos. 5,302,523 and 5,464,765, both to Zeneca. Electroporation technology has also been used to transform plants. *See* WO 87/06614 to Boyce Thompson Institute; U.S. Pat. Nos. 5,472,869 and 5,384,253, both to Dekalb; and WO 92/09696 and WO 93/21335, both to Plant Genetic Systems. Furthermore, viral vectors can also be used to produce transgenic plants expressing the protein of interest. For example, monocotyledonous plant can be transformed with a viral vector using the methods described in U.S. Pat. Nos. 5,569,597 to Mycogen Plant Science and Ciba-Giegy, now Novartis, as well as U.S. Pat. Nos. 5,589,367 and 5,316,931, both to Biosource.

[00200] As mentioned previously, the manner in which the DNA construct is introduced into the plant host is not critical to this invention. Any method which provides for efficient transformation may be employed. For example, various methods for plant cell transformation are described herein and include the use of Ti or Ri-plasmids and the like to perform *Agrobacterium* mediated transformation. In many instances, it will be desirable to have the construct used for transformation bordered on one or both sides by T-DNA borders, more specifically the right border. This is particularly useful when the construct uses *Agrobacterium tumefaciens* or *Agrobacterium rhizogenes* as a mode for transformation, although T-DNA borders may find use with other modes of transformation. Where *Agrobacterium* is used for plant cell transformation, a vector may be used which may be introduced into the host for homologous recombination with T-DNA or the Ti or Ri plasmid present in the host. Introduction of the vector may be performed via electroporation, tri-parental mating and other techniques for transforming gram-negative bacteria which are known to those skilled in the art. The manner of vector transformation into the *Agrobacterium* host is not critical to this invention. The Ti or Ri plasmid containing the T-DNA for recombination may be capable or incapable of causing gall formation, and is not critical to said invention so long as the vir genes are present in said host.

[00201] In some cases where *Agrobacterium* is used for transformation, the expression construct being within the T-DNA borders will be inserted into a broad spectrum vector such as pRK2 or derivatives thereof as described in Ditta *et al.*, (PNAS USA (1980) 77:7347-7351 and EPO 0 120 515, which are incorporated herein by reference. Included within the expression construct and the T-DNA will be one or more markers as described herein which allow for selection of transformed *Agrobacterium* and transformed plant cells. The particular marker employed is not essential to this invention, with the preferred marker depending on the host and construction used.

[00202] For transformation of plant cells using *Agrobacterium*, explants may be combined and incubated with the transformed *Agrobacterium* for sufficient time to allow transformation thereof. After transformation, the *Agrobacteria* are killed by selection with the appropriate antibiotic and plant cells are cultured with the appropriate selective medium. Once calli are formed, shoot formation can be encouraged by employing the appropriate plant hormones according to methods well known in the art of plant tissue culturing and plant regeneration. However, a callus intermediate stage is not always necessary. After shoot formation, said plant cells can be transferred to medium which encourages root formation thereby completing plant

regeneration. The plants may then be grown to seed and said seed can be used to establish future generations. Regardless of transformation technique, the gene encoding a bacterial toxin is preferably incorporated into a gene transfer vector adapted to express said gene in a plant cell by including in the vector a plant promoter regulatory element, as well as 3' non-translated transcriptional termination regions such as Nos and the like.

[00203] In addition to numerous technologies for transforming plants, the type of tissue which is contacted with the foreign genes may vary as well. Such tissue would include but would not be limited to embryogenic tissue, callus tissue types I, II, and III, hypocotyl, meristem, root tissue, tissues for expression in phloem, and the like. Almost all plant tissues may be transformed during dedifferentiation using appropriate techniques described herein.

[00204] As mentioned above, a variety of selectable markers can be used, if desired. Preference for a particular marker is at the discretion of the artisan, but any of the following selectable markers may be used along with any other gene not listed herein which could function as a selectable marker. Such selectable markers include but are not limited to aminoglycoside phosphotransferase gene of transposon Tn5 (Aph II) which encodes resistance to the antibiotics kanamycin, neomycin and G418, as well as those genes which encode for resistance or tolerance to glyphosate; hygromycin; methotrexate; phosphinothricin (bialaphos); imidazolinones, sulfonylureas and triazolopyrimidine herbicides, such as chlorsulfuron; bromoxynil, dalapon and the like.

[00205] In addition to a selectable marker, it may be desirous to use a reporter gene. In some instances a reporter gene may be used with or without a selectable marker. Reporter genes are genes that are typically not present in the recipient organism or tissue and typically encode for proteins resulting in some phenotypic change or enzymatic property. Examples of such genes are provided in K. Wising *et al.* Ann. Rev. Genetics, 22, 421 (1988). Preferred reporter genes include the beta-glucuronidase (GUS) of the *uidA* locus of *E. coli*, the chloramphenicol acetyl transferase gene from Tn9 of *E. coli*, the green fluorescent protein from the bioluminescent jellyfish *Aequorea victoria*, and the luciferase genes from firefly *Photinus pyralis*. An assay for detecting reporter gene expression may then be performed at a suitable time after said gene has been introduced into recipient cells. A preferred such assay entails the use of the gene encoding beta-glucuronidase (GUS) of the *uidA* locus of *E. coli* as described by Jefferson *et al.*, (1987 Biochem. Soc. Trans. 15, 17-19) to identify transformed cells.

[00206] In addition to plant promoter regulatory elements, promoter regulatory elements from a variety of sources can be used efficiently in plant cells to express foreign genes. For example, promoter regulatory elements of bacterial origin, such as the octopine synthase promoter, the nopaline synthase promoter, the mannopine synthase promoter; promoters of viral origin, such as the cauliflower mosaic virus (35S and 19S), 35T (which is a re-engineered 35S promoter, *see* U.S. Pat. No. 6,166,302, especially Example 7E) and the like may be used. Plant promoter regulatory elements include but are not limited to ribulose-1,6-bisphosphate (RUBP) carboxylase small subunit (ssu), beta-conglycinin promoter, beta-phaseolin promoter, ADH promoter, heat-shock promoters, and tissue specific promoters. Other elements such as matrix attachment regions, scaffold attachment regions, introns, enhancers, polyadenylation sequences and the like may be present and thus may improve the transcription efficiency or DNA integration. Such elements may or may not be necessary for DNA function, although they can provide better expression or functioning of the DNA by affecting transcription, mRNA stability, and the like. Such elements may be included in the DNA as desired to obtain optimal performance of the transformed DNA in the plant. Typical elements include but are not limited to Adh-intron 1, Adh-intron 6, the alfalfa mosaic virus coat protein leader sequence, the maize streak virus coat protein leader sequence, as well as others available to a skilled artisan. Constitutive promoter regulatory elements may also be used thereby directing continuous gene expression in all cells types and at all times (*e.g.*, actin, ubiquitin, CaMV 35S, and the like). Tissue specific promoter regulatory elements are responsible for gene expression in specific cell or tissue types, such as the leaves or seeds (*e.g.*, zein, oleosin, napin, ACP, globulin and the like) and these may also be used.

[00207] Promoter regulatory elements may also be active during a certain stage of the plant's development as well as active in plant tissues and organs. Examples of such include but are not limited to pollen-specific, embryo-specific, corn-silk-specific, cotton-fiber-specific, root-specific, seed-endosperm-specific promoter regulatory elements and the like. Under certain circumstances it may be desirable to use an inducible promoter regulatory element, which is responsible for expression of genes in response to a specific signal, such as: physical stimulus (heat shock genes), light (RUBP carboxylase), hormone (Em), metabolites, chemical, and stress. Other desirable transcription and translation elements that function in plants may be used. Numerous plant-specific gene transfer vectors are known in the art.

[00208] Standard molecular biology techniques may be used to clone and sequence the toxins described herein. Additional information may be found in Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989), Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, which is incorporated herein by reference.

[00209] Resistance Management. With increasing commercial use of insecticidal proteins in transgenic plants, one consideration is resistance management. That is, there are numerous companies using *Bacillus thuringiensis* toxins in their products, and there is concern about insects developing resistance to *B.t.* toxins. One strategy for insect resistance management would be to combine the TC toxins produced by *Xenorhabdus*, *Photorhabdus*, and the like with toxins such as *B.t.*, crystal toxins, soluble insecticidal proteins from *Bacillus* stains (*see, e.g.,* WO 98/18932 and WO 99/57282), or other insect toxins. The combinations could be formulated for a sprayable application or could be molecular combinations. Plants could be transformed with bacterial genes that produce two or more different insect toxins (*see, e.g.,* Gould, 38 *Bioscience* 26-33 (1988) and U.S. Patent No. 5,500,365; likewise, European Patent Application 0 400 246 A1 and U.S. Patents 5,866,784; 5,908,970; and 6,172,281 also describe transformation of a plant with two *B.t.* crystal toxins). Another method of producing a transgenic plant that contains more than one insect resistant gene would be to first produce two plants, with each plant containing an insect resistance gene. These plants could then be crossed using traditional plant breeding techniques to produce a plant containing more than one insect resistance gene. Thus, it should be apparent that the phrase "comprising a polynucleotide" as used herein means at least one polynucleotide (and possibly more, contiguous or not) unless specifically indicated otherwise.

[00210] Formulations and Other Delivery Systems. Formulated bait granules containing spores and/or crystals of the subject *Paenibacillus* isolate, or recombinant microbes comprising the genes obtainable from the isolate disclosed herein, can be applied to the soil. Formulated product can also be applied as a seed-coating or root treatment or total plant treatment at later stages of the crop cycle. Plant and soil treatments of cells may be employed as wettable powders, granules or dusts, by mixing with various inert materials, such as inorganic minerals (phyllosilicates, carbonates, sulfates, phosphates, and the like) or botanical materials (powdered corncobs, rice hulls, walnut shells, and the like). The formulations may include spreader-sticker adjuvants, stabilizing agents, other pesticidal additives, or surfactants. Liquid formulations may be aqueous-based or non-aqueous and employed as foams, gels, suspensions, emulsifiable concentrates, or

the like. The ingredients may include rheological agents, surfactants, emulsifiers, dispersants, or polymers.

[00211] As would be appreciated by a person skilled in the art, the pesticidal concentration will vary widely depending upon the nature of the particular formulation, particularly whether it is a concentrate or to be used directly. The pesticide will be present in at least 1% by weight and may be 100% by weight. The dry formulations will have from about 1-95% by weight of the pesticide while the liquid formulations will generally be from about 1-60% by weight of the solids in the liquid phase. The formulations will generally have from about 10^2 to about 10^4 cells/mg. These formulations will be administered at about 50 mg (liquid or dry) to 1 kg or more per hectare.

[00212] The formulations can be applied to the environment of the pest, *e.g.*, soil and foliage, by spraying, dusting, sprinkling, or the like.

[00213] Another delivery scheme is the incorporation of the genetic material of toxins into a baculovirus vector. Baculoviruses infect particular insect hosts, including those desirably targeted with the toxins. Infectious baculovirus harboring an expression construct for the toxins could be introduced into areas of insect infestation to thereby intoxicate or poison infected insects.

[00214] Insect viruses, or baculoviruses, are known to infect and adversely affect certain insects. The affect of the viruses on insects is slow, and viruses do not immediately stop the feeding of insects. Thus, viruses are not viewed as being optimal as insect pest control agents. However, combining the toxin genes into a baculovirus vector could provide an efficient way of transmitting the toxins. In addition, since different baculoviruses are specific to different insects, it may be possible to use a particular toxin to selectively target particularly damaging insect pests. A particularly useful vector for the toxins genes is the nuclear polyhedrosis virus. Transfer vectors using this virus have been described and are now the vectors of choice for transferring foreign genes into insects. The virus-toxin gene recombinant may be constructed in an orally transmissible form. Baculoviruses normally infect insect victims through the mid-gut intestinal mucosa. The toxin gene inserted behind a strong viral coat protein promoter would be expressed and should rapidly kill the infected insect.

[00215] In addition to an insect virus or baculovirus or transgenic plant delivery system for the protein toxins of the present invention, the proteins may be encapsulated using *Bacillus thuringiensis* encapsulation technology such as but not limited to U.S. Pat. Nos. 4,695,455;

4,695,462; 4,861,595 which are all incorporated herein by reference. Another delivery system for the protein toxins of the present invention is formulation of the protein into a bait matrix, which could then be used in above and below ground insect bait stations. Examples of such technology include but are not limited to PCT Patent Application WO 93/23998, which is incorporated herein by reference.

[00216] Plant RNA viral based systems can also be used to express bacterial toxin. In so doing, the gene encoding a toxin can be inserted into the coat promoter region of a suitable plant virus which will infect the host plant of interest. The toxin can then be expressed thus providing protection of the plant from insect damage. Plant RNA viral based systems are described in U.S. Pat. Nos. 5,500,360 to Mycogen Plant Sciences, Inc. and U.S. Pat. Nos. 5,316,931 and 5,589,367 to Biosource Genetics Corp.

[00217] In addition to producing a transformed plant, there are other delivery systems where it may be desirable to reengineer the bacterial gene(s). For example, a protein toxin can be constructed by fusing together a molecule attractive to insects as a food source with a toxin. After purification in the laboratory such a toxic agent with "built-in" bait could be packaged inside standard insect trap housings.

[00218] Mutants. Mutants of the DAS1529 and DB482 isolates of the invention can be made by procedures that are well known in the art. For example, an asporogenous mutant can be obtained through ethylmethane sulfonate (EMS) mutagenesis of an isolate. The mutants can be made using ultraviolet light and nitrosoguanidine by procedures well known in the art.

[00219] All patents, patent applications, provisional applications, and publications referred to or cited herein are incorporated by reference in their entirety to the extent they are not inconsistent with the explicit teachings of this specification.

[00220] Following are examples that illustrate procedures for practicing the invention. These examples should not be construed as limiting. All percentages are by weight and all solvent mixture proportions are by volume unless otherwise noted.

Example 1 – Isolation and Discovery of Insecticidal Activity of DAS1529 as a *Paenibacillus* sp.

[00221] A bacterial strain, designated DAS1529, was found to produce factors that were growth

inhibitory to neonates of lepidopteran insects, corn earworm (*Heliothis zea;* CEW), tobacco budworm (*Heliothis virescens;* TBW), and tobacco hornworm (*Manduca sexta;* THW).

[00222] DAS1529 was cultured in 2% Protease Peptone No. 3 (PP3) medium (Difco Laboratories, Detroit, MI) supplemented with 1.25% NaCl or in JB medium supplemented with 0.2% glucose. Bacterial culture was grown at 25° C for ~40 hours at 150 rpm.

[00223] The insecticidally active factors were initially found in the fermentation broth that was concentrated on 5 kDa molecular weight cutoff filters. Those factors were heat labile (inactivated by heating at 85° C for 20 minutes). These data indicated that the factors were proteinaceous. *See also* end of Example 4.

[00224] To identify active factors in cell pellets, the bacterial culture was centrifuged at 8000 rpm at 4° C for 15 minutes, washed once with sterile distilled water, and resuspended to 33X of the original culture volume in sterile distilled water, and subjected to insect bioassay as described below in Example 3. The bioassay data for DAS1529 strain is summarized in **Table 4.** The data showed that the culture broth and concentrated DAS1529 bacterial cells conferred good activity against CEW (30 to 50% mortality at 33X) and TBW (100% mortality at 33X). Those toxin factors in DAS1529 have significant relevance to the development of commercial transgenic products targeting lepidopteran insects (*e.g.* CEW and TBW) in corn and cotton.

Table 4
Bioassay of DAS 1529 Strain

Insects	TBW	CEW	Grubs
Broth Activity	+++	+++	n.d.
Pellet Activity	+++	++	-

*n.d. - not determined; -, ++, +++, no, moderate and high activity, respectively

Example 2 – Classification of DAS 1529

[00225] Molecular phylogeny was performed to determine the taxonomic affiliation of strain DAS1529. The nucleotide sequence of the 16S rDNA of DAS 1529 was determined and used for similarity and phylogenetic analyses (using the MicroSeq Kit from ABI). The sequence is provided as SEQ ID NO:16. BLAST search results are as follows:

	Core (bits)	E value
gi|15395282|emb|AJ320490.1|PTH320490 Paenibacillus thiamino...	2906	0.0

```
gi|3328014|gb|AF071859.1|AF071859     Paenibacillus popilliae s...  2834
0.0
gi|3328015|gb|AF071860.1|AF071860     Paenibacillus popilliae s...  2815
0.0
gi|2769591|emb|Y16129.1|PS16SC168     Paenibacillus sp. C-168 1...  2699
0.0
gi|2769590|emb|Y16128.1|PS16ST168     Paenibacillus sp. T-168 1...  2509
0.0
gi|2077917|dbj|D78475.1|D78475        Paenibacillus
                                      thiaminolyticu...             2503
0.0
gi|3328016|gb|AF071861.1|AF071861     Paenibacillus lentimorbus...  2493
0.0
gi|2895560|gb|AF039408.1|             Bacillus tipchiralis 16S
                                      ribosoma...                   2493
0.0
gi|2077936|dbj|D88513.1|D88513        Paenibacillus
                                      thiaminolyticu...             2493
0.0
gi|15395283|emb|AJ320491.1|PAL320491  Paenibacillus alvei pa...     2404
0.0
```

[00226] These same top scoring sequences from the BLAST search were also compared using the Gap routine (Needleman and Wunsch, J. Mol. Biol. 48; 443-453 (1970)) from GCG version 10.2, with the following results:

```
                                                                    %Ident
%Sim
gi|15395282|emb|AJ320490.1|PTH320490 Paenibacillus thiamino...     99.2  99.6
gi|3328014|gb|AF071859.1|AF071859     Paenibacillus popilliae s...  99.2
99.6
gi|3328015|gb|AF071860.1|AF071860     Paenibacillus popilliae s...  99.2
99.3
gi|2769591|emb|Y16129.1|PS16SC168     Paenibacillus sp. C-168 1...  97.1
97.3
gi|2769590|emb|Y16128.1|PS16ST168     Paenibacillus sp. T-168 1...  97.4
97.4
gi|2077917|dbj|D78475.1|D78475        Paenibacillus
                                      thiaminolyticu...             96.5
98.1
gi|3328016|gb|AF071861.1|AF071861     Paenibacillus lentimorbus...  98.8
98.9
gi|2895560|gb|AF039408.1|             Bacillus tipchiralis 16S
                                      ribosoma...                   96.0
96.9
gi|2077936|dbj|D88513.1|D88513        Paenibacillus
                                      thiaminolyticu...             96.7
98.7
gi|15395283|emb|AJ320491.1|PAL320491  Paenibacillus alvei pa...     95.2
95.3
[%Ident, matches of unambiguous bases; %Sim, %Ident plus potentially
matching ambiguous bases]
```

[00227] A number of related sequences, including the top scoring sequences noted above, were also trimmed and aligned as described by Shida *et al.* (*Int. J. Syst. Bacteriol.* 47:289-298, 1997), using the sequence alignment program CLUSTAL W (Thompson, J.D., D. G. Higgins, and T.J.

Gibson, *Nucleic Acids Res.* 22:4673-4680, 1994). The results clearly place DAS1529 in the *Paenibacillus popilliae/Paenibacillus lentimorbus* subcluster of the genus *Paenibacillus* identified by Pettersson *et al.* (*Int. J. Syst. Bacteriol.* 49:531-540, 1999), and are consistent with the analyses reported above. This subcluster includes the insect-associated species *P. popilliae* and *P. lentimorbus*, as well as *P. thiaminolyticus, Paenibacillus* sp. T-168 and C-168, and "*Bacillus tipchiralis*," which are not known to have an insect association (Pettersson *et al.*, 1999). As noted by Wayne *et al.* (*Int. J. Syst. Bacteriol.* 37:463-464, 1987) and Vandamme *et al.* (*Microbiol. Rev.* 60:407-438), rDNA sequences that are greater than 97% identical cannot generally be used to assign a bacterial strain to a particular species in the absence of additional information. In the case of DAS1529, insecticidal activity on lepidoptera and evidence of a thiaminase are not consistent with known *P. popilliae* and *P. lentimorbus*, and the insect association is not consistent with known *P. thiaminolyticus* (as well as the other subcluster species).

[00228] As other *Paenibacillus* strains are known causative agents of milky disease in larvae of Japanese beetles (*Popillia jalonica*; Harrison *et al.*, 2000), the DAS1529 was tested for activity on June beetles, a relative of Japanese beetles. No activity was found for cultures grown in JB and PP3 medium. Microscopic examination of those cultures revealed even-colored rods with no visible sporulation or parasporal crystals present. We are able to show DAS1529 can sporulate in defined medium and culture conditions and within the hemolymph of *Manducca sexta*. It is known that the Japanese beetle active *Paenibacillus* strains are typically associated with paraspore and parasporal bodies (Harrison *et al.*, 2000).

[00229] Additional work will be needed to determine whether DAS1529 belongs to an existing species or should be assigned a new species designation.

Example 3 – Insect Bioassay Methodology

[00230] Two insect bioassay methods were used to obtain results presented below. A 96-well format and a 128-well format were used for primary screening for activity against lepdidopteran insects. A 24-well diet incorporation format was used to determine specific activity (LC50s) of the toxin.

[00231] For the 96-well format, artificial diet was dispensed into 96-well microtiter plates. Each well measured approximately 0.32 cm^2 and contained 150 µl artificial diet. Samples/toxins were

applied at a rate of 50 µl /well for fermentation broth, cell pellets, and purified toxins. Positive control (*Cry*1Ac) at appropriate doses and negative controls (water, medium blank, bacterial host strains not expressing target toxin) at top dose were included. Samples were allowed to dry for approximately 1-3 hours so that the samples lost their moisture but the diet retained its moisture. Either insect eggs were dispersed onto the surface of the sample treated diet, or a single insect larva was seeded per well. The infested plate was sealed either with iron-on mylar covering or covered with sticky lidding with perforations. Tiny air holes were made in the mylar covering to ensure air supply to the insects. The plates were incubated at 28° C for 5 days and scored for mortality and stunting. This was done on a per-well basis, ignoring the number of larvae per well, as multiple eggs are often deposited per well. Activity scores were then assigned to each treatment: 0 = no activity, larvae healthy like water control wells, 1= larvae were stunted, or stunted with some mortality, 2 = larvae were all dead.

[00232] The specific activities (LC50s) of samples/toxins were determined by diet incorporation bioassay in 24-well Nutrend trays (Nu-Trend™ Container Corp., Jacksonville, FL). Insect artificial diet was made just prior to use and held in liquid state at 55° C in a water bath. Serial dilutions (≥ 5) were made by mixing 27 ml of artificial diet with no more than 3 ml of samples/toxins. A total of 30 ml sample and diet mixture was vortexed for 30 seconds and then evenly distributed into each tray, filling ~ 50% of the well volume. Trays were allowed to cool for at least 30 minutes prior to infesting. One test insect was infested into each well, and clear mylar was sealed over the top of each tray to contain the insects. Small holes were punched with an insect pin in the mylar over each well for air circulation. Assays were generally held at 25° C for 6 days but some may have been held at 30° C for 4 days if quicker results were needed. A set of positive and negative controls was run for each assay. Assays were graded on the basis of mortality but data on stunting was also recorded. Statistical methods were used to estimate LC50s for assayed samples and was expressed as ng or µg/ml diet.

Example 4 – Biochemical Purification and Characterization of Insecticidal Toxins from DAS1529 Fermentation Broth – Thiaminase

[00233] The fermentation broths of DAS1529 contained insecticidal activity against lepidopteran species, such as tobacco budworm, corn earworm, and tobacco hornworm. The nature of the insecticidal activity was investigated by biochemical purification and characterization. Corn

earworm bioassay, as described in Example 3, was used during the purification process to follow insecticidal activities.

[00234] Fermentation broths of DAS1529 were produced using 2% PP3 supplemented with 1.25% NaCl and processed as described in Example 1. Four liters of broth was concentrated using an Amicon (Beverly, MA) spiral ultrafiltration cartridge Type S1Y10 (molecular weight cut off 10 kDa) attached to an Amicon M-12 filtration device according to the manufacturer's recommendations. The retentate was diafiltered with 20 mM sodium phosphate, pH 7.0 (Buffer A) and applied at 5 ml/min to a Q cepharose XL anion exchange column (1.6 x 10 cm). The column was washed with 5 bed volumes of Buffer A to remove unbound proteins. Toxin activity was eluted by 1.0 M NaCl in Buffer A.

[00235] The fraction containing the insecticidal activity was loaded in 20 ml aliquots onto a gel filtration column Macro-Prep SE1000/40 (2.6 x 100 cm) which was equilibrated with Buffer A. The protein was eluted in Buffer A at a flow rate of 3 ml/min. Fractions with activity against corn earworm were pooled and were applied to a MonoQ (1.0 x 10 cm) column equilibrated with 20 mM Tris-HCl, pH 7.0 (Buffer B) at a flow rate of 1 ml/min. The proteins bound to the column were eluted with a linear gradient of 0 to 1 M NaCl in Buffer B at 2 ml/min for 60 min. Two milliliter fractions were collected and activity was determined as described in Example 1.

[00236] Solid $(NH_4)_2SO_4$ was added to the above active protein fractions to a final concentration of 1.7 M. Proteins were then applied to a phenyl-Superose (1.0 x 10cm) column equilibrated with 1.7 M $(NH_4)_2SO_4$ in 50 mM potassium phosphate buffer, pH 7 (Buffer C) at 1 ml/min. After washing the column with 10 milliliters of Buffer C, proteins bound to the column were eluted with a linear gradient Buffer C to 5 mM potassium phosphate, pH 7.0 at 1 ml/min for 120 min. The most active fractions determined by bioassay were pooled and dialyzed overnight against Buffer A.

[00237] The dialyzed sample was applied to a Mono Q (0.5 x 5 cm) column which was equilibrated with Buffer B at 1 ml/min. The proteins bound to the column were eluted at 1 ml/min by a linear gradient of 0 to 1 M NaCl in Buffer B. The active fractions were pooled and adjusted to a final $(NH_4)_2SO_4$ concentration of 1.7M. Proteins were then applied to a phenyl-Superose (0.5.0 x 5cm) column equilibrated with Buffer C at 1 ml/min. Proteins bound to the column were eluted with a linear gradient of Buffer C to 5 mM potassium phosphate, pH 7.0 at

0.5 ml/min for 40 min. The purified fractions were pooled and dialyzed overnight against Buffer A.

[00238] The molecular weight of the final purified toxin was examined by a gel-filtration column Superdex S-200. The toxin exhibited a native molecular weight of approximately 40 kDa. SDS-PAGE of the purified toxins showed a predominant band of approximately 40 kDa. This suggested that the native DAS1529 toxin (in this fraction) was an approximately 40 kDa monomer.

[00239] The purified toxin was electrophoresed in 4-20 % SDS-PAGE and transblotted to PVDF membrane. The blot underwent amino acid analysis and N-terminal amino acid sequencing (SEQ ID NO. 17). Searching protein database (NCBI-NR) using the sequence as a query revealed that it is 95% identical to the approximately 42 kDa thiaminase I from *Bacillus thiaminolyticus* (Campobasso *et al.,* 1998; GENBANK Accession No. 2THIA; SEQ ID NO:18). Partial sequence alignments are illustrated in **Figure 3,** which would be the same alignment with GENBANK Accession No. AAC44156 (thiaminase I precursor; U17168 is the corresponding entry in GENBANK for the DNA, which could be expressed to get a thiaminase produced and secreted from a bacterial cell). The purified thiaminase from DAS1529 was tested on corn earworm (CEW), the results were shown in **Figure 4.** This toxin did not kill corn earworm (up to a concentration of 8μg/cm^2) but exhibited 95% growth inhibition at a concentration as low as 5 ng/cm^2. It was also found that the purified thiaminase was not deactivated by proteinase K.

Example 5 – Cloning of Genes Encoding Insecticidal Factors Produced by DAS 1529

[00240] In an attempt to clone the nucleotide sequence(s) that encode the insecticidal factor(s) produced by DAS 1529, a cosmid library was constructed using total DNA isolated from DAS 1529 and was screened for insecticidal activity. Six recombinant cosmid clones were identified that produced insecticidal activity against corn earworm and tobacco budworm neonates. Three of the cosmid clones produced heat labile (when heated at 85° C for 20 minutes) factors that resulted in insect mortality. The other three cosmid clones produced heat labile factors that were growth inhibitory to insects. One of the cosmids that produced insect mortality, designated as cosmid SB12, was chosen for nucleotide sequence analysis.

[00241] A. Construction of a cosmid library of DAS1529.

[00242] Total DNA was isolated from DAS1529 with a DNA purification kit (Qiagen Inc., Valencia, CA). Vector and insert DNA preparation, ligation, and packaging, followed instructions from the supplier (Stratagene, La Jolla, CA). The DAS1529 DNA inserts as *Sau*3A I DNA fragments were cloned into the *Bam*HI site of SuperCos 1 cosmid vector. The ligated product was packaged with the Gigapack® III gold packaging extract and transfected into host cells XL1-Blue MRF'. Transformants were selected on LB-kanamycin agar plates. The cosmid library consisted of 960 randomly picked colonies that were grown in ten 96-well microtiter plates in 200 µl LB-kanamycin (50 µg/ml) for insect activity screening and long term storage.

[00243] B. Screening of DAS1529 cosmid library for insecticidal activity.

[00244] For the primary screening for clones active against lepdidopteran insects (CEW and TBW), a total of 960 cosmid clones as single colonies were grown in 2 ml cultures in 96 well plates. Cultures were spun and re-suspended in original culture media at approximately 10 X concentration and submitted to bioassay. The SuperCos 1 vector (SB1) was included as a negative control. Sixteen positive clones (SB2 to SB17) were isolated from the first round of screening. Second and third rounds of screening were carried out to screen for activity against TBW and CEW; one cosmid clone (SB12) consistently showing the highest activity was chosen for further analysis. **Table 5** summarizes the activity spectrum (as tested) of the SB12 cosmid. (BAW is beet armyworm, *Spodoptera exigua;* ECB is European cornborer, *Ostrinia nubilalis;* SCRW is Southern corn rootworm, *Diabrotica undecimpucata howardi.*) The broth of SB12 *E. coli* culture both contained no CEW activity; hence, the active factors in SB12 were different from the active factors in DAS1529 strain culture broth.

Table 5
Bioassay of SB12 *E. coli* Clone

Insects	TBW	CEW	ECB	BAW	Grubs	SCR
Broth Activity	-	-	n.d.*	n.d.	n.d.	n.d.
Pellet Activity	+++	++	+	++	-	-

*n.d. - not determined; -, ++, +++, no, moderate and high activity, respectively

C. Sequencing of SB12 cosmid insert and identification of tc- and *cry*-like ORFs.

[00245] Nucleotide sequencing of cosmid SB12 showed that it contained a genomic insert of approximately 34kb. Analysis of this sequence surprisingly revealed the presence of at least 10

putative open reading frames (ORFs) (*see* **Figure 2**). Six of the identified ORFs were surprisingly found to have some degree of amino acid sequence identity (38-48%) to *tcaA, tcaB, tcaC*, and *tccC* previously identified from *Photorhabdus luminescens* (Waterfield *et al.*, 2001), *Xenorhabdus nematophilus* (Morgan *et al.*, 2001), *Serratia entomophila* (Hurst and Glare, 2002; Hurst *et al.*, 2000), and *Yersinia pestis* (Cronin *et al.*, 2001). Those TC protein genes from *Photorhabdus, Xenorhabdus*, and *Serratia* have been shown to encode insecticidal factors. Also very interesting was that one DAS 1529 ORF had ~ 40% amino acid sequence identity to *Cry*1Ac from *Bacillus thuringiensis*, another gene previously identified as an insecticidal factor (Schnepf *et al.*, 1998; de Maagd *et al.*, 2001). Those findings have significant implication in understanding toxin gene distribution in the bacterial kingdom and in developing further strategies for toxin gene mining and engineering.

[00246] The nucleotide sequence of the SB12 cosmid was determined. The assembled DNA of 41,456 bp long was further analyzed. Three gaps were located: two in the cosmid vector and one in the insert. Analysis of the nucleotide sequence of the longest contig of approximately 34,000 bp revealed the presence of at least 10 putative open reading frames (ORFs), identified as potential start codons followed by extended open reading frames. This method is known to mis-identify translational start sites by 19% (*Bacillus subtilis*) and 22% (*Bacillus halodurans*) in genomes related to *Paenibacillus* (Besemer, J., Lomsadze, A., Borodovsky, M., *Nucleic Acids Res.* 29:2607-2618, 2001). Therefore, the quality and position of bases complementary to the *B. subtilis* 16S rRNA 5′ end (reviewed in Rocha, E.P.C., Danchin, A., Viari, A., *Nucleic Acids Res.* 27:3567-3576, 1999), N-terminal amino acid sequencing, and alignments to related genes were considered in identifying the native translation initiation sites. The putative ORFs and annotations are summarized in **Table 6** and are discussed in more detail below.

Table 6
Sequence annotation for SB12 cosmid sequence

SEQ ID NO:	Some ORF similarity to:	ORF Designation on SB12	Comments	Sequence Location on SEQ ID NO:1
1			Entire insert of SB12	(1-33521)
2	tcaA (truncated at 5')	ORF1		1-3264
3			Translation of ORF1	(1-3261)
4	tcaB	ORF2 (with IS		3271-4740 (5′ end);

		element removed)		6079-8226 (3′ end)
5			Translation of ORF2 (without insertion) from 5′-most ATG	(amino acids 1-490/ 491-1205)
6	tcaA	ORF3		9521-12820
7			Translation of ORF3	(9521-12817)
8	tcaB	ORF4		12827-16453
9			Translation of ORF4 from 5′-most ATG	(12827-16450)
10	tcaC	ORF5		16516-20850
11			Translation of ORF5	(16516-20847)
12	tccC	ORF6		20867-23659
13			Translation of ORF6 (from better RBS)	(20867-23656)
14		ORF7 (*Cry*1529)		24422-26213
15			Translation of ORF7	
19	tccC		Translation from 5′-most ATG of ORF6	20798-23656

[00247] ORF1 begins with the first nucleotide of the cloning site for the DAS1529 DNA in cosmid SB12, and is therefore missing its native translation initiation site. ORF1 shares significant DNA sequence homology with ORF3, and sequence comparison analysis suggests the first 18 bp of ORF1 is truncated, and that the first six codons encode the amino acids Met-Val-Ser-Thr-Thr, as found in OFR3. The ORF1 translation start is presumably similar to that of ORF3, from approximately bases 9505 through 9523 of SEQ ID NO:1. Two predicted amino acid sequences are presented for ORF2, ORF4, and ORF6 (SEQ ID NOs:19 and 13), based on alternative translation initiation sites, as noted above. For ORF2, SEQ ID NO:5 is discussed above. The alternate, and preferred, start site is at residue 3295 of ORF1. Thus, the protein resulting from this start site would begin at amino acid residue 9 of SEQ ID NO:5 (translation from better RBS). Likewise, for ORF4, SEQ ID NO:9 is discussed above. The alternate, and preferred, start site is at residue 12,852 of SEQ ID NO:1. The resulting protein would also be missing the first eight amino acids of SEQ ID NO:9 (thus beginning with amino acid residue 9 of SEQ ID NO:8 – translation from better RBS).

Example 6 – Sequence Analysis of "Duplicated" TCs.

[00248] The degree of sequence identity for the two ORF2 fragments ($tcaB_1$) compared to ORF4 ($tcaB_2$) was determined, as was that for ORF1 ($tcaA_1$) compared to ORF3 ($tcaA_2$). A similar

sequence relationship was observed for both pairs of ORFs.

[00249] ORF2 was constructed by combining two fragments, because of an insertion sequence-like element which was inserted in nature (apparently spontaneously), and disrupted this ORF. *See* **Figure 2.** The location of this insertion is determinable by analyzing/comparing the entire SB12 DNA sequence (SEQ ID NO:1) with the sequence of SEQ ID NO:4, the latter of which does not reflect the (non-coding) insertion. As indicated with brackets in **Figure 7,** the sequence of the 5′ translation product prior to residue 490 of SEQ ID NO:4 and the 3′ translation product from residue 491 on, align well with ORF4 (SEQ ID NO:8). The DNA sequence at the apparent insertion point shows a 9bp direct repeat commonly found flanking insertion elements (CACCGAGCA, bases 4734-4742 and 6071-6080 of SEQ ID NO:1).

Example 7 – Further Sequence Analysis

[00250] In summary, according to Vector NTI clustalW, GCG, and/or Blastp analyses, six of the identified ORFs (ORF1 to ORF6) had 38-48% amino acids sequence identity to *tcaA, tcaB, tcaC,* and *tccC* (previously identified *Photorhabdus* tc genes). The ORF7 encoded a protein that shared ~40% amino acid sequence identity to *Cry*1Ac from *Bacillus thuringiensis*, another gene previously identified as an insecticidal factor. A phylogram was generated by incorporating ORF7 (*Cry*1529) sequence with a large number of other *Cry* proteins (**Figure 8**). This phylogenetic tree suggests that *Cry*1529 is distantly related to other *P. popilliae Cry* sequences such as the *Cry*18s (Zhang *et al.,* 1997, Zhang *et al.,*1998) that are closer to *Cry*2s; *Cry*1529 falls (remotely but most closely) into a group of *Cry* proteins containing commonly found lepidoptera (*Cry*1, *Cry*9), coleoptera (*Cry*3, *Cry*8, *Cry*7), and diptera (*Cry*4) toxins, which is a distinct group compared to those including nematode toxins *Cry*5, -12, -13, -14, and -21 and *Cry*2, -18.

[00251] It was a surprising and novel discovery to find *Cry* and TC protein genes (in the SB12 genomic insert) in *Paenibacillus*. The identification of new *Cry* and TC protein genes has relevance to the art's understanding of *Photorhabdus* and *Xenorhabdus,* and *Bacillus thuringiensis*, and broadens the scope of bacterial genera in which *Cry* and TC protein genes have been found. The size of the full-length *Cry*1529 identified herein corresponds to the core toxin of *Cry*1s; *Cry*1529 represents a new class of *Cry* proteins which also has implications for isolating further *cry* genes from *Bacillus thuringiensis* and *Paenibacillus*.

[00252] To verify that these surprising observations were not the result of strain contamination

(*i.e.,* to confirm that the 34 kb insert carrying TC and *Cry* ORFs was indeed from the total DNA of DAS1529), molecular analysis was carried out by Southern blot hybridization and PCR. For PCR validation, ORF6 (*tccC*-like) and ORF7 (*Cry*1529) specific primers (Example 8, **Table 8**) were used to amplify ORF6 and ORF7 from SB12 cosmid and DAS1529 total DNA. For ORF6, PCR amplifications were performed on a PE9600 thermal cycler (Perkin Elmer) with the following parameters: initial denaturation at 95° C for 2 minutes; 30 cycles each with denaturing at 95° C for 30 seconds, annealing at 60° C for 45 seconds, extension at 72° C for 2 minutes, and a final extension for 10 minutes at 72° C. For ORF7, amplification parameters were the same as ORF6, except the annealing temperature was 55° C for 30 seconds and extension at 72° C for 4 minutes. Specific PCR products with a single band of expected sizes were amplified for both ORF6 and ORF7.

[00253] Initial southern blot hybridization was based on partial SB12 DNA sequence and carried out according to standard protocol (Sambrock *et al.,* 1990). DNA samples included total DNA of DAS1529 from two independent preparations, SB12 cosmid DNA, and one negative control DNA sample from NC1 (*Photorhabdus*). Both DAS1529 DNA samples were 16S rDNA sequence confirmed to be of *Paenibacillus sp.* origin, and one was originally used for cosmid library construction; the other was a new preparation. DNA samples were digested with *Eco*RI, blotted onto membrane, and hybridized with Roche DIG System (Roche) labeled 180 bp of PCR product amplified out of SB12. The PCR primers are 5' CCT CAC TAA AGG GAT CAC ACG G 3' annealing partially to the vector and truncated ORF1 (compared to full-length ORF3), and 5' GGC TAA TTG ATG AAT CTC CTT CGC 3' annealing to the truncated ORF1 (*tcaA*-like) and full length ORF3 (*tcaA*-like). A total of three DNA fragments (0.85, 2.7, and 8.0 kb) hybridizing to the PCR probe were detected, 0.85 and 8.0 in the SB12 and 2.7 and 8.0 in DAS1529 DNAs. No signals were detected in the negative control. The 0.85 kb (from first *Eco*RI ORF1 internal fragment to first *Eco*RI site in the vector) and 8.0 kb (from first 5' *Eco*RI site in ORF3 to the third *Eco*RI site in ORF1) matched the calculated sizes of the target DNA fragments from SB12. Detection of the 2.7 kb fragment suggests the presence of an *Eco*R1 site 2.7 kb immediately upstream of the first *Eco*RI site within ORF1 in DAS1529 DNA. Those results show that the SB12 insert was from DAS1529 total DNA and, based on hybridization and restriction analysis, all copies of the ORFs were accounted for.

Example 8 – Characterization of Insecticidal Activities of Proteins Encoded by SB12 Cosmid ORFs

[00254] Random transposon insertional mutagenesis (to disrupt an individual ORF or an entire operon) and heterologous expression (expressing individual ORFs or entire operons) were undertaken to isolate individual ORF(s) or operons conferring the insecticidal activities in the SB12 cosmid.

[00255] A. Random transposon mutagenesis of SB12 cosmid

[00256] A Tn mutagenesis library was generated from DAS1529 cosmid SB12 using the GPS-1 Genome Priming System (New England BioLabs, Beverly, MA) following the kit instructions. Briefly, 2μl 10X GPS buffer, 1μl pGPS2.1 Donor DNA (0.02 μg), 1μl SB12 cosmid (0.1 μg) and 18 μl sterile H_2O were added to a 0.5 ml tube. One μl of TnsABC Transposase was added; the mixture was vortexed and then spun briefly to collect the materials at the bottom of the tube. This reaction mixture was incubated for 10 minutes at 37° C. One μl of Start Solution was added and mixed by pipetting up and down several times. The reaction was incubated at 37° C for 1 hour and was then heat inactivated at 75° C for 10 minutes. One μl of the reaction mixture was diluted 10-fold with sterile H_2O; 1 μl of the diluted reaction was electroporated into 100 μl of Electro MAX DH5α-E *E. coli* (Gibco BRL, Rockville, MD). After 1 hour of outgrowth in SOC medium at 37° C, 10 μl or 100 μl were plated on LB agar plates containing 20 μg/ml Kanamycin and 15 μg/ml chloramphenicol, followed by incubation overnight at 37 °C.

[00257] Individual colonies from the SB12 Tn mutagenesis were streaked onto fresh LB agar plates containing 20 μg/ml Kanamycin and 15 μg/ml chloramphenicol. From the streaks, 50 ml cultures of LB containing 20 μg/ml Kanamycin and 15 μg/ml chloramphenicol were inoculated and grown at 28°C, 200 rpm for 48 hours. The cells were then collected by centrifugation at 3500 rpm for 20 minutes. The supernatant was removed and the pellet resuspended in 2.5 ml of the culture supernatant for a 20X concentration. The concentrated cell pellet was then assayed for activity against corn earworm. Forty μl of the 20X concentrate was surface applied to artificial diet using 8 wells per sample in 128 well plates. Newly hatched corn earworm larvae were added and allowed to feed for 5 days, at which time mortality and weights were recorded.

[00258] A total of 184 clones were tested for loss of activity against corn earworm. The results are summarized in **Table 7.** Bioassay of Tn clones revealed that a Tn insertion in the *Cry*1529 gene results in complete loss of activity. Initial bioassay showed that the activities of clones

which carried Tn insertions in the tc genes were variable. Further analysis of those clones in which cultures were all normalized to the same cell density prior to bioassay showed no loss of activity as compared to SB12. Results from Tn analysis suggest that ORF7(*Cry*1529) is the key insecticidally active component of SB12 cosmid.

Table 7
Bioassay of SB12, *Cry*1529 and tc tn insertion *E. coli* Clones

Insects	TBW	CEW	THW	Grubs	SCRW
SB12	+++	++	+++	-	n.d
Tn in *Cry*1529	-	-	-	-	-
Tn in tcs	+++	++	+++	-	-

* n.d. - not determined; -, ++, +++, no, moderate and high activity, respectively

[00259] B. Heterologous expression SB12 ORFs/operon.

[00260] *Cry*1529 (ORF7) and five tc ORFs (*see* **Table 8** below) were expressed in pET101D® system. *See* **Figure 5.** This expression vector has all the attributes of the basic T7-regulated pET expression system (Dubendorff and Studier, 1991; Studier and Moffatt, 1986) and allows directional cloning of a blunt-end PCR product into a vector for high-level, regulated expression, and simplified protein purification in *E. coli.* Optimal PCR amplification employed high-fidelity *PfuTurbo*™ DNA polymerase that is highly thermostable and possesses a 3' to 5' exonuclease proofreading activity to correct nucleotide-misincorportaion errors (Stratagene, La Jolla, CA). When ThermalAce™ polymerase (Invitrogen) is used, point mutations were introduced in the tc ORFs, which were corrected by the *PfuTurbo*™ based Quick-Change™ XL site-directed mutagenesis kit (Stratagene). The *E. coli* strain BL21 Star™ (DE3), was used as a host for expression since it contains the *rne131* mutation (Lopez *et al.,* 1999) that generally enhances mRNA stability and the yield of the recombinant proteins.

[00261] Individual ORFs were PCR amplified out of the SB12 cosmid with ORF specific primers (**Table 8**) under defined conditions. As a directional cloning requirement, the forward PCR primers were designed to contain the sequence, CACC, at the 5' end to ensure PCR product base pair with the overhang sequence, GTGG, in the pET101.D vector. The reverse primers when paired with forward primers will amplify each ORF, respectively. PCR reactions were carried out in 50 µl reaction mixture containing of 50 ng of SB12 cosmid DNA, 1X *Pfu* reaction buffer (Stratagene), 0.2mM each of dNPT, 0.25 mM of each primer, and 2 U of *PfuTurbo* DNA

polymerase (Stratagene). PCR amplifications were performed on a PE9600 thermal cycler (Perkin Elmer) with the following parameters: initial denaturation at 95° C for 2 minutes, 35 cycles each with denaturing at 95° C for 30 seconds, annealing at 55° C for 30 seconds, extension at 72° C for 2 minutes per kb ORF, and a final extension for 10 minutes at 72° C.

Table 8.
Summary of PCR Primers for Cloning ORF1-7

ORFs	Forward primers (5' to 3')	Reverse primers (5' to 3')
ORF1 ($tcaA_1$)	CACCATGCTTTATAAGGCCTGGC	TCAGGCCTGCACCGC
ORF3 ($tcaA_2$)	CACCATGGTGTCAACAACAGACAACAC	TCAGGCTTTCGCTGCAGC
ORF4 ($tcaB_2$)	CACCATGACCAAGGAAGGTGATAAGC	CTATTTCATAACATATCGAATTGG
ORF5 (tcaC)	CACCATGCCACAATCTAGCAATGC	TCACCGCGCAGGCGGTGAAG
ORF6 (tccC)	CACCATGAAAATGATACCATGGACTCATC	CTACTTTCTCTTCATTGAAAACCGGCGG
ORF7 (*Cry*1529)	CACCATGAACTCAAATGAACCAAATTTATC	AACTGGAATTAACTTCGATTC

[00262] PCR products for each ORF were cloned into pET101.D following instructions from the supplier (Invitrogen). The cloned ORF was purified as pET101.D plasmid DNA and sequenced verified. Since Tn analysis indicated ORF7 is the key component of SB12 for control of the tested pests, biochemical analysis and insect bioassay focused on heterologously expressed ORF7 proteins. For ORF7 expression clones, DNA sequence analysis showed 100% match with the original SB12 DNA sequence. Expression of ORF7 was induced by 0.5 mM IPTG for 4 hrs according to kit instructions (Invitrogen).

[00263] C. Bioassay for insecticidal activities of ORF7 and tc operon.

[00264] Bioassay samples were prepared as whole *E. coli* cells, cell lysates, and purified toxins. The spectrum and specific activity of ORF7 (*Cry*1529) is summarized in **Table 10.** *Cry*1529 is most active against tobacco hornworm (*Manduca sexta*) and highly active (LC50 of 10 μg/ml diet) against tobacco budworm (*Heliothis virescens*); 100% mortality was observed for both insects. At higher doses, *Cry*1529 conferred some mortality (20 to 60%) and substantial growth inhibition on corn earworm (*Heliothis zea*), beet armyworm (*Spodoptera exigua*), and black cutworm (*Agrotis ipsilon*). For European cornborer (*Ostrinia nubilalis*), *Cry*1529 had some growth inhibition at higher doses. For some other insect species (fall armyworm, boll weevil,

southern rootworm, mosquito), no activity was detected. The *Cry*1529 LC50s for *Cry*1A (*Cry*1Ac) resistant diamond back moth (DBMr) and sensitive diamond back moth (DBM) colonies are >50 μg/ml and <1.0 μg/ml, respectively, suggesting a potential cross resistance. *Cry*1529 did not confer detectable activity on grass grubs, a relative of Japanese beetles.

[00265] To test the activity of other non-*Cry*1529 factors in DAS1529, one *Cry*1529 tn knockout SB12 cosmid clone (tn67) was assayed against TBW, CEW, SCRW, ECB, BW, BAW, THW, and grass grubs; no activity was found against these pests. To address the issue of potential non- or low-expression of tc ORFs in SB12 background, individually expressed tc ORFs were tested independently and in combination with the other TCs from DAS 1529; no activity was found against TBW, CEW, and grass-grubs. Further, four ORFs were expressed as a single operon to very high levels in *E. coli* cells. When tested *in vitro,* the whole cells contained no detectable activity on TBW, CEW, and grass-grubs. While the lack of grub activity is somewhat interesting, these results are not surprising in that *Paenibacillus* typically infect a narrow range of grub hosts. In light of this, it could follow that the spectrum of activity of the insecticidal toxins might also be relatively narrow. Thus, screens (using known methods) involving a broader range of pests, and additional time, would be required to identify susceptible pests. The results presented herein should not lead one away from recognizing that the subject TC proteins have utility as do other TC proteins from *Xenorhabdus*, *Photorhabdus*, and the like.

[00266] Soluble proteins were extracted with 25 mM sodium phosphate pH 8.0, 100 mM sodium chloride and analyzed on 4-12% NuPAGE gradient gel with 1X MES buffer (Invitrogen). ORF7 protein was purified using standard procedures, and N-terminal sequencing revealed the expected sequence: MNSNEPNLSDV. A bioassay was performed with whole *E. coli* cells, with normalized cell density, expressing target proteins. *See* **Figure 6.** Large scale purified ORF7 protein was used to obtain LC50s for ORF7 by *in vitro* bioassay. Thermal stability analysis of the purified ORF7 indicated that a 5 minute treatment at 75° C is sufficient to abolish its activity against TBW. *See* **Table 9.**

Table 9.
Thermal Stability of Purified *Cry*1529 (ORF7)

Samples	Activity
*Cry*1529, room temperature	+++
*Cry*1529, 50° C for 5 min.	+++
*Cry*1529, 50° C for 10 min.	+++
*Cry*1529, 75° C for 5 min.	-
*Cry*1529, 75° C for 10 min.	-
*Cry*1529, 100° C for 5 min.	-

-, +++, no and full activity, respectively

[00267] For the tc genes, error-free clones of ORF3 and ORF6 were used as intermediate clones to generate a tc operon clone expressing ORF3 (*tcaA*), ORF4 (*tcaB*), ORF5 (*tcaC*), and ORF6 (*tccC*). To construct the tc operon in pET101.D, the *NsiI*/*Sac*I fragment containing partial *tcaA*, entire *tcaB* and *tcaC*, and partial *tccC* was excised out of SB12 cosmid to replace the *NsiI*/*Sac*I insert in pET101.D-tcaA; this was followed by the insertion of a 208 bp *Sac*I fragment from pET101.D-tccC, *See* **Figure 5.** All four ORFS were expressed to high levels by standard IPTG induction. For the ORF6 (tccC) expressed in the tc operon, the size of the expressed protein was slightly smaller than the ORF6 predicted by Vector NTI from the 5'-most ATG (SEQ ID NO:18) and expressed independently. Hence, the annotated ORF6 (SEQ ID NO:13) based on the presence of a ribosome binding site consensus is likely the native protein produced in SB12 and DAS1529.

D. Activity Spectrums of Toxins

[00268] The toxin activity spectrum of *Cry*1529 (ORF7) is summarized in **Table 10.**

Table 10.
Spectrum activity for *E. coli* and *Pseudomonas* expressed *Cry*1529

Species	Active (+++)	Format & Method	Material Production Method	LC_{50}
H. virescens (TBW)	+++	96 well top load and diet incorporation (scores, mortality, inhibition)	FCP, SE, purified, IC	11μg tox/ml diet with *E. coli* cell preps
H. zea (CEW)	+	96 well top load and diet incorporation (scores, mortality, inhibition)	FCP, SE, purifed, IC	>100 μg tox/g diet
S. exigua (BAW)	+	96 well top (score)	FCP, purifed	>78 $\mu g/cm^2$
S. frugiperda (FAW)	-	96 well top (score)	FCP, purifed	>>10$\mu g/cm^2$
Plutella xylostella (DBM)	+++	96 well top (score)	FCP, purifed	0.02 μg tox/g diet
*Cry*1A resistant *Plutella xylostella* (rDBM)	+	96 well top (score)	FCP, purifed	59.7μg tox/g diet

A. ipsilon (BCW)	+	96 well top (score)	FCP, purifed	>10 $\mu g/cm^2$
O. nubilalis (ECB)	+	128 well top (weights)	FCP, purifed	>43 $\mu g/cm^2$
Culex sp. (Mosquito)	-	1oz cups (mortality)	FCP, purifed	>20 $\mu g/ml$ H_20
Diabrotica undecimpunctata howardi (SCRW)	-	96 well top (score)	FCP, purifed	>>100 μg tox/cm^2
Anthonomous grandis grandis (BW)	-	128 well top (weights)	FCP, purifed	>>43 μg tox/cm^2
M. sexta (THW)	+++			(highly active)
Continis mutabilis (Beetles); surrogate for grass grub	-			>>100 μg tox/g soil

Key: -, +, ++, +++ (no, low, moderate, high activity); FCP, frozen cell pellets; SE, soluble extract; purified = column purified *Cry*1529; IC, P.f. *Cry*1529 inclusion

[00269] Only a limited range of pests was used in assays in an initial attempt to determine the activity spectrum of the subject TCs/tc ORFs. The following data, using the ORF3-OR6 operon, were obtained:

Table 11.
Spectrum activity for Tc ORF's

Species	Active (+++)	Format & Method	High Dose	Material Production Method	Comments
H. virescens (TBW)	-	96 well top (score)	10x	FCP	No effect
H. zea (CEW)	-	96 well top (score)	10x	FCP	No effect
S. exigua (BAW)	-	96 well top (score)	10x	FCP	No effect
Spodoptera frugiperda (FAW)	-	96 well top (score)	10x	FCP	No effect
A. ipsilon (BCW)	-	96 well top (score)	10x	FCP	No effect

[00270] Again, while this initial round of screening did not reveal activity of these TCs against these pests, one skilled in the art would not doubt that the subject proteins are useful, as are the corresponding *Photorhabdus/Xenorhabdus* proteins. In addition, *see* Example 10, below.

Example 9 – Use of PCR primers for identifying *Cry*1529 homologues from other bacterial genera, species, and strains

[00271] For screening additional ORF7 *cry*1529 homologs from other (*Paenibacillus* or other) strains, gene specific and degenerate PCR primers were designed to amplify the target ORF7 DNA sequences of 1 kb. The PCR primers were deduced from two, well-conserved protein

motifs (QAANLHL, domain I, block 1 core for forward primer; GPGFTGGD, domain III, block 3 for reverse primer) highly conserved in *Cry* proteins. Those primers are listed in **Table 12** and were validated on DAS1529. PCR amplifications were performed on a PE9600 thermal cycler (Perkin Elmer) with the following parameters: initial denaturation at 95° C for 2 minutes; 35 cycles each with denaturing at 95° C for 30 seconds, annealing at 47° C for 45 seconds, extension at 72° C for 2 minutes, and a final extension for 10 minutes at 72° C. Those primer pairs were used to screen a bacterial (non-*B. thuringiensis*) culture collection by PCR. Five out of 192 strains (three *Paenibacillus*, one *Bacillus*, and one unidentified) produced PCR products of expected sizes. These strains were also found to have CEW activity according to primary bioassay screening. However, sequence analysis of amplicons obtained from one of these strains using different primers showed that the amplicons were not derived from a *cry* gene.

[00272] Notwithstanding this, and as these screens were not exhaustive, the subject invention includes methods of screening *Paenibacillus* spp., *Bacillus* spp. (including *Bacillus thuringiensis* and *sphaericus*), and the like for *Cry*1529-like proteins and genes. Given the significant nature of the discovery of lepidopteran-toxic *Cry* proteins in *Paenibacillus*, the subject invention also includes methods of screening *Paenibacillus* spp., generally, for lepidopteran-toxic *Cry* proteins and genes. Various screening methods are well-known in the art, including PCR techniques (as exemplified above), probes, and feeding assays (where whole cells are fed to target pests). As one skilled in the art would readily recognize, screening methods of the subject invention include the preparation and use of clone libraries (such as cosmid libraries) in these screens.

Table 12
PCR Primers for Screening ORF7 Homologs

Gene-specific and degenerate Primers	DNA sequence (5' to 3')
*Cry*1529-F	CAAGCAGCCAACCTCCACCTA
*Cry*1529-R	ATCCCCTCCTGTAAAGCCTGG
*Cry*PP-F	CAAGCNGCNAATYTWCATYT
*Cry*PP-R	TCNCCNCCNGTAAANCCWGG
*Cry*PT-F	CARGCSGCSAAYYTBCAYYT
*Cry*PP-F2	CAAGCWGCWAATYTWCATYT
*Cry*PP-R2	TCHCCWCCWGTAAAWCCWGG
*Cry*PT-F2	CAGGCSGCSAAYYTGCATYT

1529=gene specific; PP=*P. popilliae* codon bias; PT=*P. thiaminolyticus* codon bias

Example 10 – Complementation of *Xenorhabdus* XptA2 TC Protein Toxin with DAS1529 TC Proteins

[00273] This example provides experimental evidence of the ability of DAS1529TC proteins, expressed here with a single operon (ORFs 3-6; tcaA, tcaB, TcaC and tccC; *see* section C of Example 8), to complement the XptA2 toxin from *Xenorhabdus nematophilus* Xwi (*see* SEQ ID NO:49). Two independent experiments were carried out to express the DAS1529 TC operon and XptA2 independently, or to co-express the XptA2 gene and the TC operon in the same *E. coli* cells. Whole cells expressing different toxins/toxin combinations were tested for activity against the lepidopteran insects: corn earworm (*Heliothis zea; CEW*) and tobacco budworm (*Heliothis virescens; TBW*). The data from both experiments indicate that DAS1529 TC proteins are able to enhance *Xenorhabdus* TC protein XptA2 activity on both insect species tested.

[00274] A. Co-expression of DAS1529 TCs and *Xenorhabdus* XptA2

[00275] Expression of the TC operon was regulated by the T7 promoter/*lac* operator in the pET101.D expression vector that carries a ColE1 replication origin and an ampicillin resistance selection marker (Invitrogen). Comprehensive description of cloning and expression of the tc operon can be found in section C of Example 8. The XptA2 gene was cloned in the pCot-3 expression vector, which carries a chloramphenicol resistance selection marker and a replication origin compatible with the ColE1. The pCot-3 vector expression system is also regulated by the T7 promoter/*lac* operator. Hence, compatible replication origins and different selection markers form the basis for co-expression of the TC operon and XptA2 in the same *E. coli* cells. Plasmid DNAs carrying the TC operon and XptA2 were transformed into *E. coli*, BL21 Star™ (DE3) either independently or in combination. Transformants were selected on LB agar plates containing 50 μg/ml carbenicillin for pET101.D-TC operon, 50μg/ml chloramphenicol for pCot-3-XptA2, and both antibiotics for pET101.D-TC operon/pCot-3-XptA2. To suppress basal toxin expression, glucose at a final concentration of 50 mM were included in both agar and liquid LB medium.

[00276] For toxin production, 5ml and 50 ml of LB medium containing antibiotics and 50 mM glucose were inoculated with overnight cultures growing on the LB agar plates. Cultures were grown at 30°C on a shaker at 300 rpm. Once the culture density has reached an O.D. of ~ 0.4 at 600 nm, IPTG at a final concentration of 75 μM was added to the culture medium to induce gene expression. After 24 hours, *E. coli* cells were harvested for protein gel analysis by the NuPAGE

system (Invitrogen). Cell pellets from 0.5 ml 1X culture broth was resuspended in 100 µl of 1X NuPAGE LDS sample buffer. Following brief sonication and boiling for 5 min, 5 µl of the sample was loaded onto 4 to 12% NuPAGE bis-tris gradient gel for total protein profile analysis. XptA2 expressed to detectable levels when expressed independently or in the presence of the TC operon. Based on gel scan analysis by a Personal Densitometer SI (Molecular Dynamics), XptA2 expressed nearly 8X as high by itself as when co-expressed with the TC operon. For the 5 ml induction experiment, there is a nearly equal expression of XptA2.

[00277] B. Bioassay for Insecticidal Activity

[00278] As described in Example 8, DAS1529 tc ORFs when expressed independently or as an operon, did not appear to be active against TBW and CEW. The following bioassay experiments focused on determining whether *Paenibacillus* (DAS1529) TC proteins (of ORFs 3-6; TcaA-, TcaB-, TcaC-, and TccC-like proteins) can complement *Xenorhabdus* TC protein toxin activity (XptA2 is exemplified). Bioassay samples were prepared as whole *E. coli* cells in 4 X cell concentrate for the 5ml induction experiment, both the XptA2 and XptA2/TC operon cells contained very low but nearly equal amount of the XptA2 toxin. Data in Table 13 showed that at the 4X cell concentration tested, TC proteins + *Xenorhabdus XptA2* was active against CEW. This provided the first evidence of a complementation effect of *Paenibacillus* DAS1529 TC proteins on *Xenorhabdus XptA2.*

Table 13 Bioassay of DAS1529 TC complementation of *Xeno.* XptA2 on *H. zea*

Insects:	CEW
Negative control	-
TCs (DAS1529)	-
Xeno. XptA2	-
TC proteins + *Xeno.* XptA2	++

* -, ++, +++ = no, moderate and high activity, respectively

[00279] For the second bioassay experiment, the amount of XptA2 protein in the XptA2 cells and the XptA2 + TC operon cells was normalized based on densitometer gel scan analysis. As shown in Table 14, XptA2 *per se* had moderate activity at 40X on TBW (*H. virescens*), but the activity dropped to a level undetectable at and below 20X. However, when co-expressed withTCs, high levels of activity were very apparent in the presence of 10X and 5X XptA2, and low activity was still noticeable at 1.25X XptA2. These observations indicate there is a significant potentiation

effect of 1529 TC proteins on *Xenorhabdus* XptA2 against *H. virescens*. At the highest doses tested, neither the negative control nor the tc operon *per se* had any activity against this pest .

Table 14 Bioassay of IDAS1529 TC complementation of XptA2 on *H. virescens*

Normalized XptA2	40X	20X	10X	5X	2.5X	1.25X
XptA2	+	-	-	-	n.d.	n.d.
TCs +XptA2	n.d.	n.d.	++	++	+	-

* n.d. - not determined; -, +, ++, +++ = no, low, moderate, and high activity, respectively

Example 11 – Stabilization of *Cry*1529 protein against trypsin digestion.

[00280] This example teaches modifications to the DNA sequence disclosed as SEQ ID NO:14, which encodes the *Cry*1529 protein (disclosed as SEQ ID NO:15) such that the new encoded proteins are more resistant to proteolytic digestion by trypsin than is the native protein. Digestion of proteins in the gut of insects limits the time of exposure of the insect to a protein toxin. Therefore, methods that decrease the susceptibility of a protein toxin to protease digestion can be used to increase potency of the protein.

[00281] For these tests, trypsin enzyme (*e.g.* Sigma Chemical #T1426) and trypsin inhibitors (*e.g.* Sigma Chemical #T9008) were prepared as stock solutions of 4 mg/mL or 10 mg/mL in 50 mM Tris HCl buffer, pH8.0. Test incubations with various concentrations of trypsin and *Cry*1529 protein were performed at 37°C for 1 hour, and were terminated by addition of an equal volume of an equal concentration of trypsin inhibitors (*e.g.* a digestion that received 35 µL of 4 mg/mL trypsin solution was terminated by addition of 35 µl of 4 mg/mL trypsin inhibitors). For a typical experiment, *Cry*1529 protein was produced by appropriately engineered *E. coli* cells and purified by steps described previously, which included separation from other proteins by passage through a size-exclusion column. Following digestion, the protease products were analyzed by standard acrylamide gel electrophoresis followed by immunoblot analysis using antibody prepared against the *Cry*1529 protein. The results of such an experiment are shown in **Figure 9.**

[00282] Trypsin digestion produces two major protein products, the smaller of which is approximately 50 kDa in molecular size. It is noted that this digestion pattern is the same as that produced from trypsin digestion of a *Cry*1529-His_6 protein, which is identical to the native *Cry*1529 protein amino acid sequence of SEQ ID NO:15 except for the addition of amino acids KGELNSKLEGKPIPNPLLGLDSTRTGHHHHHH to the carboxy-terminus. The coding region

for *Cry*1529-His_6 was produced by ligating the coding region for the native *Cry*1529 protein into the pET101/D-TOPO® vector (Invitrogen™, Carlsbad, CA). This recombinant clone was made to facilitate purification of the recombinant *Cry*1529 protein by binding to a commercially available V5 antibody, whose epitope is represented by the amino acid sequence GKPIPNPLLGLDSTRTG (underlined above), or by purification schemes that expoit the six histidine residues (double underlined above). Procedures for these manipulations were performed according to the recommendations provided with the pET101/D-TOPO® vector.

[00283] Trypsin digestion of the *Cry*1529-His_6 protein was found to eliminate activity in insect bioassays against lepidopteran insects. MALDI-TOF analysis was used to determine the sequence of amino acids composing the N-terminus of the 50 kDa peptides, and two protease processing sites were determined, corresponding to amino acid residues 122 (R, Arginine) and 126 (K, Lysine) of SEQ ID NO:15.

[00284] Modifications to remove the first trypsin cleavage site in the encoded protein were made in the native DNA sequence (SEQ ID NO:14), using the QuickChange® mutagenesis methodology (Stratagene, La Jolla, CA). Three different types of mutations were made at amino acids in the region of 120 to 123 of SEQ ID NO:15: RARA to HANA, RARA to RARS, and RARA to QANA. The DNA oligonucleotide primers (listed in the 5' to 3' direction for each strand) for these mutations are listed in Table 15 below. The bases that differ from the native DNA sequence are underlined.

Table 15

Mutation	Forward (Coding strand) Primer	Reverse (Complementary strand) Primer
RARA to HANA (pMYC2865)	AAAATGATTCTAATAATTTACACGCGAAC GCTGTAGTGAAAGAC	GTCTTTCACTACAGCGTTCGCGTGTAAATTA TTAGAATCATTTT
RARA to QANA (pMYC2866)	AAAATGATTCTAATAATTTACAAGCGAAC GCTGTAGTGAAAGAC	GTCTTTCACTACAGCGTTCGCTTGTAAATTA TTAGAATCATTTT
RARA to RARS (pMYC2867)	AAAATGATTCTAATAATTTAAGAGCGAGA TCTGTAGTGAAAGAC	GTCTTTCACTACAGATCTCGCTCTTAAATTA TTAGAATCATTTT

[00285] Comparison of the wild type and mutated coding regions induced by these primers are shown in this Table. The pertinent amino acid residues are shown in bold type.

Table 16

Wild-type:	gAA	AAT	GAT	TCT	AAT	AAT	TTA	**AGA**	GCG	**AGA**	**GCT**	GTA	GTG	AAA	GAC
Amino Acids:	(E)	N	D	S	N	N	L	**R**	A	**R**	**A**	V	V	K	D
			115					120		122	123		125	126	
RARA to HANA:	gAA	AAT	GAT	TCT	AAT	AAT	TTA	**CAC**	GCG	**AAC**	GCT	GTA	GTG	AAA	GAC
Amino Acids:	(E)	N	D	S	N	N	L	**H**	A	**N**	A	V	V	K	D
RARA to QANA:	gAA	AAT	GAT	TCT	AAT	AAT	TTA	**CAA**	GCG	**AAC**	GCT	GTA	GTG	AAA	GAC
AMINO ACIDS:	(E)	N	D	S	N	N	L	**Q**	A	**N**	A	V	V	K	D
RARA to RARS:	gAA	AAT	GAT	TCT	AAT	AAT	TTA	AGA	GCG	AGA	**TCT**	GTA	GTG	AAA	GAC
Amino Acids:	(E)	N	D	S	N	N	L	R	A	R	**S**	V	V	K	D

[00286] The separate, mutated coding regions were each cloned into the pET101/D-TOPO® vector, which allows inducible production of the *Cry*1529 variant proteins. *E. coli* cells containing the constructs were grown, and expression of the *Cry*1529 variant genes was induced by methods recommended by the supplier. Harvested whole cells were then tested in trypsin digestion assays, and analyzed as above. Typical results are shown in **Figure 10.** For these experiments, 10 mg of whole cell pellet was suspended in 50 mM Tris HCl, pH8.0, and digested for 3 hours at 37° in a final volume of 1 mL, with 100 µL of 10 mg/mL trypsin. The reactions were mixed occasionally during incubation. Digestion was terminated by addition of 100 µL of 10 mg/mL trypsin inhibitors and the tubes were stored on ice.

[00287] These results demonstrate that both the native *Cry*1529 (RARA) and the *Cry*1529-His_6 (RARA) proteins are digested by trypsin to produce a major product of about 50 kDa. When the RARA sequence corresponding to the trypsin cleavage site was mutated to HANA or QANA, substantial resistance to trypsin digestion was obtained. No 50 kDa peptides were produced, and easily detectable amounts of the apparently full-length *Cry*1529-His_6 proteins were present. Mutation of the RARA site to RARS did not eliminate production of the 50 kDa peptides, but substantially reduced the rate of protease cleavage. Thus, it is apparent that mutation of protease processing sites in the *Cry*1529 protein substantially decreases its susceptibility to protease digestion. This allows the proteins to reside for longer periods of time in the insect gut following ingestion, resulting in increased potency to kill susceptible insects.

Example 12 – Design of PCR Primers for Detection of homologues of IDAS 1529 tcORFs in other *Paenibacillus* strains

[00288] As shown above, *Paenibacillus* strain IDAS 1529 produces an extracellular protein that is toxic to various Lepidopteran insects. Molecular phylogeny of the 16S ribosomal gene of this strain indicates that it is most closely related to members of the *P. thiaminolyticus-P. lentimorbus-P. popilliae* cluster. It has also been shown that *Paenibacillus* strain IDAS 1529 contains both toxin complex genes (hereafter designated as *tc* genes) and a novel insecticidal crystalline inclusion protein gene designated *cry*1529. In an attempt to determine if *tc* homologues are present in other members of the genus *Paenibacillus,* a collection of *Paenibacillus* strains was screened by polymerase chain reaction (PCR) and hybridization analyses. For the PCR analyses, total DNA isolated from *Paenibacillus* strains was used as template and screened using oligonucleotide primers specific to *tc* genes found in *Paenibacillus* strain IDAS 1529, *Photorhabdus* species, and *Xenorhabdus* species. Amplified products obtained with the *tc* primer sets were cloned and their nucleotide sequence was determined and compared to *tc* sequences obtained from *Paenibacillus* strain IDAS 1529. The following Examples illustrate how one can design *tc*-specific oligonucleotide primers and use PCR to search the total DNA of *Paenibacillus* isolates for DNA sequences that are homologous to *tc* genes identified in *Paenibacillus* strain IDAS 1529, *Photorhabdus* species, and *Xenorhabdus* species. By using PCR analysis (as described herein), it was (and is) possible to identify *tc* homologues in a species of *Paenibacillus* distinct from *Paenibacillus* strain IDAS 1529 and the *P. thiaminolyticus-P. lentimorbus-P. popilliae* cluster.

12.A. – Extraction of total DNA from *Paenibacillus* strains

[00289] *Paenibacillus* strains were grown on nutrient agar plates (8 g/l nutrient broth, 15 g/l Bacto agar; Difco Laboratories, Detroit, MI) for 3-5 days at 30°C. A single colony was picked and inoculated into a 500 ml tribaffled flask containing 100 ml of sterile nutrient broth (8 g/l nutrient broth; Difco Laboratories, Detroit, MI). Following 24-72 hrs of incubation at 30°C on a rotary shaker at 150 rpm, the cultures were dispensed into sterile 500 ml polyethylene bottles and centrifuged at 6,500xg for 1 hr at 4°C. After centrifugation, the supernatant fluid was decanted and the bacterial cell pellet was retained. Total DNA was extracted from the cell pellet using the QIAGEN Genomic-tip System 100/G and associated Genomic DNA Bufffer Set (QIAGEN Inc., Valencia, CA, USA) by following The Sample Preparation and Lysis Protocol for Bacteria exactly as described by the manufacturer. The extracted total DNA was solubilized in 0.5 ml TE buffer (10 mM Tris-HCl, pH 8.0; 1 mM EDTA, pH 8.0).

12.B. – Selection of *tc* specific oligonucleotide primers for PCR

[00290] To select oligonucleotide primers specific to the *tc* genes previously identified from *Paenibacillus* strain IDAS 1529, the *tcaA*, *tcaB*, *tcdB* and *tccC* nucleotide sequences obtained from *Paenibacillus* strain IDAS 1529, *Photorhabdus* strain W14, and *Xenorhabdus* strain Xwi were aligned using the Align program in the Vector NTI software package (Informax, Inc., Frederick, MD). Nucleotide sequences used for this analysis are listed in **Table 17**.

Table 17. Nucleotide sequences used for *tc* specific primer selection

Gene	Source organism	Source of nucleotide sequence	Gene Designation
tcaA1	*Paenibacillus* strain IDAS 1529	SEQ ID NO:2	*tcaA1-1529*
tcaA2	*Paenibacillus* strain IDAS 1529	SEQ ID NO:6	*tcaA2-1529*
tcaA	*Photorhabdus* strain W14	GenBank: Accession No. AF346497	*tcaA-W14*
tcaB1	*Paenibacillus* strain IDAS 1529	SEQ ID NO:4	*tcaB1-1529*
tcaB2	*Paenibacillus* strain IDAS 1529	SEQ ID NO:8	*tcaB2-1529*
tcaB	*Photorhabdus* strain W14	GenBank: Accession No. AF346497	*tcaB-W14*
tcdB1	*Photorhabdus* strain W14	SEQ ID NO:42	*tcdB1-W14*
tcdB2	*Photorhabdus* strain W14	SEQ ID NO:43	*tcdB2-W14*
xptC1	*Xenorhabdus* strain Xwi	SEQ ID NO:20	*xptC1-Xwi*
tcaC	*Paenibacillus* strain IDAS 1529	SEQ ID NO:10	*tcaC-1529*
tccC1	*Photorhabdus* strain W14	SEQ ID NO:44	*tccC1-W14*
tccC2	*Photorhabdus* strain W14	SEQ ID NO:45	*tccC2-W14*
tccC3	*Photorhabdus* strain W14	SEQ ID NO:46	*tccC3-W14*
tccC4	*Photorhabdus* strain W14	SEQ ID NO:47	*tccC4-W14*
tccC5	*Photorhabdus* strain W14	SEQ ID NO:48	*tccC5-W14*
xptB1	*Xenorhabdus* strain Xwi	SEQ ID NO:21	*xptB1-Xwi*
tccC	*Paenibacillus* strain IDAS 1529	SEQ ID NO:19	*tccC-1529*

12.B.i. – *tcaA* specific primer selection

[00291] Nucleotide sequence alignment of *tcaA1-1529*, *tcaA2-1529*, and *tcaA-W14* identified two regions of nucleotide sequence identity of sufficient length for the selection of PCR primers with minimal degeneracy (shown as boxed regions in **Figure 10**.). These two regions were selected for the synthesis of *tcaA* specific primers, which were designated SB105 and SB106 (**Tables 18** and **19**).

12.B.ii. – *tcaB* specific primer selection

[00292] Nucleotide sequence alignment of *tcaB1-1529*, *tcaB2-1529*, and *tcaB-W14* identified four regions of nucleotide sequence identity of sufficient length for the selection of PCR primers with

minimal degeneracy (**Figure 11**.). These four regions were selected for the synthesis of *tcaB* specific primers, which were designated as SB101, SB102, SB 103, and SB104 (**Tables 18** and **19**).

12.B.iii. – *tcaC* specific primer selection

[00293] Nucleotide sequence alignment of *tcdB1-W14, tcdB2-W14, xptC1-Xwi* and *tcaC-1529* identified two regions of nucleotide sequence identity of sufficient length for the selection of PCR primers with minimal degeneracy (**Figure 12**.). These two regions were selected for the synthesis of *tcaC* specific primers, which were designated as SB215 and SB217 (**Tables 18** and **19**).

12.B.iv. – *tccC* specific primer selection

[00294] Nucleotide sequence alignment of *tccC1-W14, tccC2-W14, tccC3-W14, tccC4-W14, tccC5-W14, xptB1-Xwi* and *tccC-1529* identified two regions of nucleotide sequence identity of sufficient length for the selection of PCR primers with minimal degeneracy (**Figure 13**.). These two regions were selected for the synthesis of *tccC* specific primers, which were designated as SB212 and SB213 (**Tables 18** and **19**).

Table 18. *tc* specific primers

Primer designation	Primer length	Sequence of primer (5' to 3')	SEQ ID NO.
SB101	32	GCKATGGCSGACCCGATGCAWTACAAGCTGGC*	22
SB102	32	AGCGGYTGACCRTCCAGRCTCARATTGTGGCG	23
SB103	28	TGTATAACTGGATGGCYGGWCGTCTSTC	24
SB104	26	TCRAAAGGCAGRAAMCGGCTGTCGTT	25
SB105	28	CTTCYCTKGATATCYTKYTGGATGTGCT	26
SB106	30	ACGRCTGGYATTGGYAATCAGCCARTCCAA	27
SB212	27	CGYTATIAATATGAYCCKGTVGGYAAT	28
SB213	25	CATCBCGYTCTTTRCCIGARTARCG	29
SB215	33	CGHAGCTCYICCCAGTWYTGGCTGGATGARAAA	30
SB217	32	GTRTCATTTTCATCTTCRTTBACIRYAAACCA	31

* K= G or T; S= G or C; W= A or T; Y= C or T; R= G or A; M= A or C; V= G or C or A; B= G or T or C; H= A or C or T, I= inosine

Table 19. *tc* primer combinations

Target gene	Forward primer	Reverse primer	Approximate size of expected amplified product
tcaA	SB105	SB106	1.4 kb
tcaB	SB101	SB102	0.4 kb
tcaB	SB103	SB104	0.65 kb

tcdB	SB215	SB217	2.2 kb
tccC	SB212	SB213	0.9 kb

Example 13 – PCR amplification of *Paenibacillus* DNA

[00295] For PCR amplification using *tcaA*- and *tcaB*-specific primer sets, 3-5 ul of total DNA obtained from each of the *Paenibacillus* strains was mixed with 50 pmoles of each primer and 1X Eppendorf MasterMix (Eppendorf AG; Hamburg, Germany) in a 20 ul reaction volume. Amplification conditions were denaturation at 94°C for 3 minutes followed by 30 cycles of denaturation at 94°C for 1 minute, annealing at 52°C for 1.5 minutes, and extension at 72°C for 1.5 minutes, followed by a final extension at 72°C for 5 minutes.

[00296] For PCR amplification using *tcaC*- and *tccC*-specific primer sets, approximately 375 ng of total DNA obtained from each of the *Paenibacillus* strains was mixed with 50 pmoles of each primer and 12.5 ul of Epicentre® FailSafe™ Buffer D and 2.5 U of Epicentre® FailSafe™ Polymerase (Epicentre; Madison, WI) in a 25 ul reaction volume. Amplification conditions were denaturation at 96°C for 4 minutes followed by 40 cycles of denaturation at 94°C for 30 seconds, annealing at 64°C for 30 seconds, and extension at 70°C for 30 seconds. Each cycle, the annealing temperature was lowered by 0.5 °C and the extension time was increased by 5 seconds.

13.A. – Gel electrophoresis, cloning, and nucleotide sequence determination of PCR amplified products

[00297] PCR amplification reactions were examined by gel electrophoresis using 0.8 to 1% Seakem LE agarose (BioWhittaker Molecular Applications, Rockland, ME) in 1X TAE buffer. Amplified products were cloned in the vector pCR 2.1-TOPO® using the TOPO TA® Cloning Kit (Invitrogen™ Life Technologies, Carlsbad, CA) exactly as described by the manufacturer. The nucleotide sequences of the cloned amplified products were determined using M13 Forward, M13 Reverse, and *tc* sequence-specific sequencing primers as needed to obtain double stranded sequence of each cloned amplified product. Nucleotide sequencing was performed using the CEQ Dye Terminator Cycle Sequencing Quick Start Kit (Beckman Coulter, Fullerton, CA, USA) and the CEQ 2000 XL DNA Analysis System (Beckman Coulter) exactly as directed by the manufacturer. The Sequencher (v. 4.1.4) software package (Gene Codes, Ann Arbor, MI) was used to construct contigs from the sequencing data and determine a consensus sequence for each amplified product.

13.B. – Nucleotide sequence analysis of PCR amplified products

13.B.i. – *tcaA*

[00298] When PCR using the *tcaA*- (primer combination SB105 and SB106) was performed using total DNA obtained from the collection of *Paenibacillus* strains, it was observed that total DNA from a *Paenibacillus apiarius* strain (NRRL NRS 1438; hereafter designated as DB482) produced an amplified product of the expected sizes. The amplified product was cloned and sequenced.

[00299] The amplified product obtained using the SB105 and SB106 primer combination was designated as *tcaA2-DB482*. When the sequence of *tcaA2-DB482* (SEQ ID NO:32) as compared to the *tcaA* sequences obtained from *Paenibacillus* strain IDAS 1529 and *Photorhabdus* strain W14, it was observed that *tcaA2-DB482* have the greatest nucleotide sequence identity (90.5% over 1,239 nucleotides) to *tcaA2-1529* (**Table 20**). The deduced amino acid sequence encoded by *tcaA2-DB482* (designated as TcaA2-DB482; SEQ ID NO:33) was 89.1% identical to the corresponding deduced amino acid sequence of *tcaA2-1529* (designated as TcaA2-1529; SEQ ID NO:7).

Table 20. Nucleotide and deduced amino acid sequence identity of *tcaA2-DB482* with corresponding regions of *tcaA1-1529, tcaA2-1529*, and *tcaA-W14*

Gene	% Nucleotide identity with tcaA2-DB482	% deduced amino acid sequence identity with tcaA2-DB482
tcaA1-1529	57	33
tcaA2-1529	90	89
tcaA-W14	50	32

13.B.ii. – tcaB

[00300] The amplified products obtained using the SB101 and SB102 primer combination and the SB103 and SB104 primer combination were designated as *tcaB2a-DB482* and *tcaB2b-DB482*, respectively. When the sequences of *tcaB2a-DB482* (SEQ ID NO:34) and *tcaB2b-DB482* (SEQ ID NO:35) were compared to the *tcaB* sequences obtained from *Paenibacillus* strain IDAS 1529 and *Photorhabdus* strain W14, it was observed that both of these sequences have the greatest nucleotide sequence identity to *tcaB1-1529* and *tcaB2-1529* (**Table 21**). The nucleotide sequence identity of *tcaB2a-DB482* and *tcaB2b-DB482* to *tcaB2-1529* was 92.6% and 89.8%, respectively. The deduced amino acid sequences encoded by *tcaB2a-DB482* (designated as

TcaB2a-DB482; SEQ ID NO:36) *tcaB2b-DB482* (designated as TcaB2b-DB482; SEQ ID NO:37) were 91.2% and 91.1% identical, respectively, to the corresponding deduced amino acid sequence of *tcaB2-1529* (designated as TcaB2-1529; SEQ ID NO:9).

Table 21. Nucleotide and deduced amino acid sequence identity of tcaB2a-DB482 and tcaB2b-DB482 with corresponding regions of tcaB1-1529, tcaB2-1529, and tcaB-W14

Gene	% Nucleotide identity with *tcaB2a-DB482*	% Nucleotide identity with *tcaB2b-DB482*	% deduced amino acid sequence with TcaB2a-DB482	% deduced amino acid sequence with TcaB2b-DB482
tcaB1-1529	93	93	94	92
tcaB2-1529	93	90	91	92
tcaB-W14	63	57	59	57

13.B.iii. – *tcdB*

[00301] When PCR using the *tcaC*-specific primer combination (SB215 and SB217) was performed using total DNA obtained from DB482 produced an amplified product of the expected size. The amplified product was cloned and sequenced.

[00302] The amplified product obtained using the SB215 and SB217 primer combination was designated as *tcaC-DB482*. When the sequence of *tcaC-DB482* (SEQ ID NO:38) was compared to the *tcaC* sequences obtained from *Paenibacillus* strain IDAS 1529, *Xenorhabdus* strain Xwi and *Photorhabdus* strain W14, it was observed that *tcaC-DB482* has the greatest nucleotide sequence identity (93.5% over 2,091 nucleotides) to *tcaC-1529* (**Table 22**). The deduced amino acid sequence encoded by *tcaC-DB482* (designated as TcaC-DB482; SEQ ID NO:39) was 91.1% identical to the corresponding deduced amino acid sequence of *tcaC-1529* (designated as TcaC-1529; SEQ ID NO:11).

Table 22. Nucleotide and deduced amino acid sequence identity of *tcaC-DB482* corresponding regions of *xptC1-Xwi, tcdB1-W14*, and *tcdB2-W14*, and *tcaC-1529*

Gene	% Nucleotide sequence identity with *tcaC-DB482*	% deduced amino acid sequence identity with TcaC-DB482
tcaC-1529	93	91
xptC1-Xwi	50	35
tcdB1-W14	50	36
tcdB2-W14	50	36

13.B.iv. – tccC

[00303] When PCR using the *tccC*-specific primer combination (SB212 and SB212) was performed using total DNA obtained from the collection of *Paenibacillus* strains, it was observed that total DNA from DB482 produced an amplified product of the expected size. The amplified product was cloned and sequenced.

[00304] The amplified product obtained using the SB212 and SB213 primer combination was designated as *tccC-DB482*. When the sequence of *tccC-DB482* (SEQ ID NO:40) was compared to the *tccC* sequences obtained from *Paenibacillus* strain IDAS 1529, *Xenorhabdus* strain Xwi and *Photorhabdus* strain W14, it was observed that *tccC-DB482* has the greatest nucleotide sequence identity (93.7% over 858 nucleotides) to *tccC-1529* (**Table 23**). The deduced amino acid sequence encoded by *tccC-DB482* (designated as TccC-DB482; SEQ ID NO:41) was 95.5% identical to the corresponding deduced amino acid sequence of *tccC-1529* (designated as TccC-1529; SEQ ID NO:13).

Table 23. Nucleotide and deduced amino acid sequence identity of *tccC-DB482* corresponding regions of *xptB1-Xwi, tc-W14, tccC-1529*, and *tcc* genes from *Photorhabdus* strain W14

Gene	% Nucleotide sequence identity with tccC-DB482	% deduced amino acid sequence identity with TccC-DB482
tccC-1529	94	96
xptB1-Xwi	54	45
tccC1-W14	54	48
tccC2-W14	56	45
tccC3-W14	56	46
tccC4-W14	56	46
tccC5-W14	54	44

13.C. – Summary of PCR analyses

[00305] This example (and other examples herein) illustrate methods for designing oligonucleotide primers based on *tc* genes from three genera of bacteria, and that the use of these primers for PCR screening of *Paenibacillus* strains can identify *tc* homologues present in those strains. DB482, which is an isolate of *Paenibacillus apiarius* (deposited as NRRL B-30670) that was isolated from honey bee larva, was shown to contain homologues of *tcaA*, *tcaB*, *tcaC*, and *tccC*. The finding of these tc homologues confirms that *Paenibacillus* strain IDAS 1529 is not unique within the genus *Paenibacillus* with regard to possessing *tc* genes. Therefore, one skilled in the art can now use these and other methods to identify other *tc* homologues in other species of

Paenibacillus such as *P. chondroitinus*, *P. alginolyticus*, *P. larvae*, *P. validus*, *P. gordonae*, *P. alvei*, *P. lentimorbus*, *P. popilliae*, *P. thiaminolyticus*, *P. curdlanolyticus*, *P. kobensis*, *P. glucanolyticus*, *P. lautus*, *P. chibensis*, *P. macquariensis*, *P. azotofixans*, *P. peoriae*, *P. polymyxa*, *P. illinoisensis*, *P. amylolyticus*, *P. pabuli*, and *P. macerans*.

Example 14 – Detection of homologues of IDAS 1529 tcORFS in other *Paenibacillus* strains by Southern Hybridization

[00306] This example illustrates how one can use radioactively labeled DNA fragments as probes to search the genomic DNA of *Paenibacillus* isolates for DNA sequences (preferably having some homology to the known tcORFs first detected in IDAS 1529). The results demonstrate that sequences homologous to two of the tcORFs are detected in a *Paenibacillus apairius* isolate, DB482.

[00307] Genomic DNA from various *Paenibacillus* strains (or from *E. coli* to serve as a negative control) was prepared as described above in Example 12, and was digested with restriction enzyme to produce multiple fragments. A typical digestion contained 8 µg of DNA in a total volume of 400 µL of reaction buffer as supplied by the manufacturer of the *Eco*R I enzyme (New England Biolabs, Beverly, MA). The reaction, containing 200 units of enzyme, was incubated overnight at 37°C, then placed on ice. Digested DNA was further purified and concentrated by addition of 30 µL of 3M sodium acetate (pH5.2) and 750 µL of ice cold 100% ethanol, followed by centrifugation. The DNA pellet was washed twice with 70% ethanol, dried under vacuum, and resuspended in 50 µl of TE buffer [10mM Tris HCl, pH8.0; 1mM ethylenediaminetetraacetic acid (EDTA)]. An aliquot was then analyzed by agarose gel electrophoresis for visual assurance of limit digestion. In a similar manner, DNA of IDAS 1529 cosmid SB12 was digested with *Eco*R I, and was used as a positive control for the hybridization experiments.

[00308] *Eco*R I digested genomic DNA fragments to be blotted for Southern analysis were separated by electrophoresis through 0.7% or 1.2 % agarose gels in TEA buffer (40mM Tris-acetate, 2mM EDTA, pH8.0) (1 µg DNA/well). On each gel, lanes containing a 1kb DNA Molecular Weight Ladder (Invitrogen™, Carlsbad, CA) were used to provide molecular weight size standards. The 15 fragment sizes larger than 500 bp in this ladder (in kilobases) are: 12.2, 11.2, 10.1, 9.2, 8.1, 7.1, 6.1, 5.1, 4.1, 3.1, 2.0, 1.6, 1.0, 0.52, and 0.50. The DNA in the gel was stained with 50 µg/mL ethidium bromide, the gel was photographed, and then the DNA in the gel

was depurinated (5 min in 0.2M HCl), denatured (15 min in 0.5M NaOH, 1.5M NaCl), neutralized (5 min in 0.2M HCl) and transferred to MAGNA 0.45 micron nylon transfer membrane (Osmonics, Westborough, MA) in 2X SSC (20X SSC contains 3M NaCl, 0.3M sodium citrate, pH 7.0). The DNA was crosslinked to the membrane by ultraviolet light (Stratalinker®; Stratagene, La Jolla, CA) and prepared for hybridization by incubating at 60°C or 65°C for 1 to 3 hours in "Minimal Hybridization" solution [contains 10% w/v polyethylene glycol (M.W. approx. 8000), 7% w/v sodium dodecylsulfate; 0.6X SSC, 5mM EDTA, 100 μg/ml denatured salmon sperm DNA, and 10mM sodium phosphate buffer (from a 1M stock containing 95 g/L $NaH_2PO_4 \cdot 1H_2O$ and 84.5 g/L $Na_2HPO_4 \cdot 7H_2O$)].

[00309] DNA fragments of the tcORFs for use as hybridization probes were first prepared by Polymerase Chain Reaction (PCR) using SB12 cosmid DNA as template (see previous examples). The forward and reverse primers for these amplifications are listed (5' to 3' directions of the respective DNA strands) in Table 24, below (bases in capital letters correspond to protein coding regions). Primer Set One is designed to amplify, from SB12 cosmid DNA, a DNA fragment that includes all of tcORF5, which is disclosed as SEQ ID NO:10, and which has some similarity to the *Photorhabdus tcaC* gene (Table 6). Primer Set Two is designed to amplify, from cosmid SB12, a DNA fragment that encodes the protein disclosed as SEQ ID NO:19. This DNA fragment and the encoded protein are somewhat longer than the DNA sequence of tcORF6 disclosed as SEQ ID NO:12, and the encoded protein disclosed as SEQ ID NO:13. The proteins disclosed as SEQ ID NO:13 and SEQ ID NO:19 both have some similarity to the protein encoded by the *Photorhabdus tccC* gene (Table 6). The amplified PCR products were cloned into the pCR®2.1-TOPO® cloning vector (Invitrogen™, Carlsbad, CA), and fragments containing the tcORFs were released from the resulting clones by restriction enzyme digestion (listed in the Table below), followed by purification from agarose gels using the GenElute™ Agarose Spin columns (Sigma Chemical Co, St Louis, MO). Recovered fragments were concentrated by precipitation using the Quick-Precip™ Plus Solution according to the supplier's instructions (Edge BioSystems, Gaithersburg, MD).

Table 24.

PCR Primer Set One
SB12 tcORF5 (SEQ ID No. 10)
Forward Primer SB126*

gtacgtcatctagaaaggagatataccATGCCACAATCTAGCAATGCCGATATCAAGCTATTGTC
Reverse Primer SB127*
tgacatcggtcgacattattaCCGCGCAGGCGGTGAAGCAAATAATGATGAGTCCATGGTA
Cut from pCR®2.1-TOPO® clone with *Sal* I + *Xba* I + *Pvu* I and purify 4,368 bp fragment

PCR Primer Set Two
SB12 tcORF that encodes SEQ ID No. 19; encompassing tcORF6 (SEQ ID No. 12)
Forward Primer SB128*, ** gtacgtcaactagtaaggagatataccATGAAAATGATACC**g**TGGACTCA**c**CATTATTTGCTTCACC **Reverse Primer SB129*** tgacatcgctcgagattattaCTTTCTCTTCATTGAAAACCGGCGGAAAAAGTTCCCA Cut from pCR®2.1-TOPO® clone with *Eco*R I + *Sph* I /+ *Pvu* I and purify 2,925 bp fragment

* In this table, bases in lower case at the 5' ends of the primers are not complementary to the cosmid SB12 DNA sequence. They were used to provide restriction enzyme recognition sequences on the ends of the amplified products to facilitate subsequent cloning manipulations.
**Bases in lowercase bold were changed from those of the native sequence to eliminate a potential hairpin structure that might interfere with subsequent functional analysis of the clone.

[00310] Radioactively labeled DNA fragments were prepared using the High Prime Radioactive Labeling Kit (Roche Diagnostics, Mannheim, Germany) according to the supplier's instructions. Nonincorporated nucleotides were removed by passage through a QIAquick® PCR Purification column (Qiagen, Inc. Valencia, CA) according to the manufacturer's instructions. Labeling of approximately 100 ng of DNA fragments by these methods resulted in specific activities of approximately 0.1 μCi/ng. The labeled DNA fragments were denatured by boiling for 5 minutes, then added to the hybridization blot in Minimal Hybridization solution and incubated overnight at 60°C or 65° C. Loose radioactivity was removed from the blot by rinsing at room temperature in 2X SSC, then more tightly bound radioactivity was removed by washing the blot for at least one hour at 60°C or 65° C in 0.3X SSC + 0.1% sodium dodecylsulfate. At least two such washes were performed. The blot was placed on X-ray film at -80°C with two intensifying screens, and the exposed film was developed after 1 to 3 days exposure. Blots were stripped of hybridized DNA fragments by boiling for 10 minutes in 0.3 X SSC + 0.1% SDS, and reused once or twice for subsequent hybridizations.

[00311] Distinct fragments that hybridized to probes derived from Primer Sets One and Two were observed in genomic DNA obtained from *Paenibacillus apairius* strain DB482. The probe derived from Primer Set One (primers SB126 and SB127), which detects sequences homologous to the IDAS 1529 tcORF5, hybridized to fragments of estimated sizes (in kilobases) of 20, 10.2, and 8.4. Within this range of molecule sizes, mobilities of DNA fragments can provide only

estimations of true molecular sizes. Signal intensity for the fragments estimated to be 20 kb and 8.4 kb were much more intense than the signal intensity for the fragment estimated to be 10.2 kb. Since each of these fragments is at least twice the size of the probe fragment (about 4.4 kb), one explanation for these results is that multiple copies of genes that are similar to the probe derived from IDAS1529 tc ORF5, and thus are similar to the *Photorhabdus tcaC* gene, are present in the genome of *Paenibacillus apairius* strain DB482. However, other explanations for this outcome are possible.

[00312] The probe derived from Primer Set Two (primers SB128 and SB129), which detects sequences homologous to the IDAS 1529 tcORF6 and its flanking 5' end sequences, hybridized to fragments of estimated sizes (in kilobases) of 7.8 and 4.5. Signal intensity for the fragment estimated to be 7.8 kb was very much more intense than the signal intensity seen for the fragment estimated to be 4.5 kb. One explanation for this result is that *Paenibacillus apairius* strain DB482 has a single gene similar to the IDAS 1529 tcORF6 and its 5' flanking sequences, and thus is similar to the *Photorhabdus tccC* gene, and that *Eco*R I cleaves the gene into two fragments that have unequal portions of the DNA sequences comprising the gene. However, other explanations for this outcome are possible, including the presence of multiple genes with different amounts of absolute homology to the probe.

[00313] These results (detection by PCR amplification followed by DNA sequence analyses) confirm the presence of relatives of the *Photorhabdus tcaC* and *tccC* genes in *Paenibacillus apairius* strain DB482.

Example 15 - Insecticidal activity of DB482

[00314] *Paenibacillus* strain DAS1529 has been shown to produce an extracellular protein that is toxic to Lepidopteran insects and has also been shown to contain a *cry* gene, designated as *cry1529*. As this strain produces an extracellular insecticidally active protein and intracellular insecticidally active proteins, the subject invention includes screening other strains of *Paenibacillus* for extracellular (released into culture supernatant fluid) and/or intacellular (cell-associated) insecticidally active agents. This example illustrates how one can produce fermentation broths of *Paenibacillus* strains, how to process these broths, and how to test samples derived from these broths for insecticidal activity.

15.A. Production and processing of *Paenibacillus* fermentation broths

[00315] *Paenibacillus* strains were grown on nutrient agar plates (8 g/l nutrient broth, 15 g/l Bacto agar; Difco Laboratories, Detroit, MI) for 3-5 days at 30°C. A single colony was picked and inoculated into a 500 ml tribaffled flask containing 100 ml of sterile modified tryptic soy broth (tryptone 10- g/l, peptone 7 g/l, soytone 3 g/l, KCl 5 g/l, K_2PO_4 2.5 g/l; Difco Laboratories, Detroit, MI). Following 72 hours of incubation at 28°C on a rotary shaker at 150 rpm, the cultures were dispensed into sterile 500 ml polyethylene bottles and centrifuged at 4,000xg for 45 minutes at 4°C. After centrifugation, the supernatant fluid was decanted and filtered through a 0.22 um membrane filter (Millipore Corporation, Bedford, MA). The culture filtrate was then concentrated 20X using a Centricon Plus-20 centrifugal filter device with a 5,000 molecular weight cutoff membrane by centrifuging at 4,000xg. The bacterial cell pellet was resuspended in 10 mM potassium phosphate buffer (pH=8). These samples were then tested in insect bioassay for insecticidal activities contained in the processed supernatant and cell pellet samples.

15.B. Insect bioassay of processed supernatant and cell pellets

[00316] The insect species included in these assays were *Diabrotica undecimpunctata howardi* (Southern corn rootworm, SCR), *Helicoverpa zea* (corn earworm, CEW), and *Heliothis virescens* (tobacco budworm, TBW) The artificial diet used to rear and bioassay SCR was described previously (Rose, R.L. and McCabe, J.M. 1973. J. Econ. Entomol. 66, 398-400). Standard artificial lepidopteran diet (Stoneville Yellow diet) was used to rear and bioassay ECB, CEW, and TBW. Forty ul aliquots of the concentrated supernatant or cell pellet samples were applied directly to the surface of wells (~1.5 cm^2) containing the artificial diet. Treated diet wells were allowed to air-dry in a sterile flow-hood, and each well was infested with a single, neonate insect hatched from surface-sterilized eggs. Assay trays were then sealed, placed in a humidified growth chamber, and maintained at 28°C for 3-5 days. Mortality and larval weight determinations were then scored. Eight insects were used per treatment.

15.C. Insecticidal activity of DB482

[00317] Concentrated supernatant and cell pellets from strain DB482 had insecticidal activity against SCR, TBW, and CEW relative to control treatments (**Table 25**.) It is possible that the insecticidal activity associated with concentrated supernatants and cell pellets from DB482 are

the result of two different insecticidal factors, one that is cell-associated (*i.e. Cry*-like) and another that is released from the cells (*i.e.* TC-like). However, it is also possible that the insecticidal activities from both the concentrated supernatant and cell pellets from DB482 are the result of the same insecticidal factors being present in both cellular fractions.

Table 25. Insecticidal activity of DB482

Insects Tested	Concentrated Supernatant activity	Cell pellet activity
SCR	+++*	+++
TBW	++	++
CEW	+++	++
Medium controls	-	-

* -, ++, +++; no, moderate, and high activity, respectively

15.D. Summary of Insecticidal Activity Screening

[00318] This example illustrates a method for screening concentrated culture supernatants and cell pellets from *Paenibacillus* strains to identify strains possessing insecticidal activity against Coleopteran and Lepidopteran insects. DB482, which is an isolate of *Paenibacillus apiarius* was shown herein to contain homologues of *tcaA*, *tcaB*, *tcaC*, and *tccC*. The finding of insecticidal activity in DB482 confirms that *Paenibacillus* strain DAS1529 is not unique within the genus *Paenibacillus* with regard to producing insecticidal activities against Lepidopteran insects. Therefore, the subject invention includes methods used to identify other strains of *Paenibacillus* with insecticidal activities against Lepidopteran insects in other species of *Paenibacillus* such as *P. chondroitinus*, *P. alginolyticus*, *P. larvae*, *P. validus*, *P. gordonae*, *P. alvei*, *P. lentimorbus*, *P. popilliae*, *P. thiaminolyticus*, *P. curdlanolyticus*, *P. kobensis*, *P. glucanolyticus*, *P. lautus*, *P. chibensis*, *P. macquariensis*, *P. azotofixans*, *P. peoriae*, *P. polymyxa*, *P. illinoisensis*, *P. amylolyticus*, *P. pabuli*, *P. macerans*.